FACULTAD DE CIENCIAS

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

PROGRAMA DE DOCTORADO DE BIOCIENCIAS MOLECULARES

CRISPR/Cas9 system and fluorescence enhanced BA-LIFT bioprinting technique as a toolbox for the study of the Immune System

TESIS DOCTORAL

MIGUEL GÓMEZ FONTELA

Madrid, 2020

FACULTAD DE CIENCIAS

DEPARTAMENTO DE BIOLOGÍA MOLECULAR

PROGRAMA DE DOCTORADO DE BIOCIENCIAS MOLECULARES

CRISPR/Cas9 system and fluorescence enhanced BA-LIFT bioprinting technique as a toolbox for the study of the Immune System

Miguel Gómez Fontela

Graduado en Biología Sanitaria

por la Universidad de Alcalá de Henares

Dirigida por:

Pilar Lauzurica Gómez y Carlos Molpeceres Álvarez

Lugar de realización:

Instituto de Salud Carlos III / Universidad Politécnica de Madrid.

Madrid, 2020

La doctora Pilar Lauzurica Gómez, investigadora científica y líder del grupo de Presentación y Regulación Inmunes del Centro Nacional de Microbiología, Instituto de Salud Carlos III, y el Doctor Carlos Molpeceres Álvarez, catedrático de la Universidad Politécnica de Madrid y director del Centro Láser UPM,

CERTIFICAN

Que la Tesis Doctoral 'CRISPR/Cas9 system and fluorescence enhanced BA-LIFT bioprinting technique as a toolbox for the study of the Immune System' ha sido realizada mediante colaboración del Instituto de Salud Carlos III y la Universidad Politécnica de Madrid, bajo su tutela, y que reúne las condiciones para optar al grado de Doctor.

Madrid, febrero 2020

Fdo. Pilar Lauzurica Gómez y Carlos Molpeceres Álvarez.

Esta tesis ha sido realizada como fruto de la colaboración entre el laboratorio de Presentación y Regulación Inmunes del Centro Nacional de Microbiología adscrito al Instituto de Salud Carlos III y el Centro Láser de la Universidad Politécnica de Madrid.
Asimismo, no hubiera sido posible sin la financiación del Ministerio de Educación, Cultura y Deporte a través de una beca de Formación del Profesorado Universitario (FPU).

Run rabbit, run.
Dig that hole, forget the sun. When at last, the work is done, don´t sit down, it's time to dig another one…

Breathe
Pink Floyd

Lo bello del desierto es que en algún lugar esconde un pozo.

El principito
Antoine de Saint-Exupéry

El destino conduce al hombre, pero es el hombre el que lo sigue porque quiere y es libre de no querer seguirlo.

Vida y destino
Vasili Grossman

El fracaso fortifica a los fuertes.

Vuelo nocturno
Antoine de Saint-Exupéry

Aspiró una bocanada de frío húmedo y el corazón se le llenó de esperanza y de luz: tal vez aquél terrible sueño había acabado.

Vida y destino
Vasili Grossman

Agradecimientos

Esta tesis no hubiera sido posible sin el inquebrantable apoyo de mi familia. Siempre han estado y siempre estarán detrás de todas las buenas decisiones que pueda tomar en la vida. Su experiencia ha cimentado las baldosas que me han permitido caminar con decisión; su altruismo me ha otorgado oportunidades que otros no han podido disfrutar; su perseverancia y apoyo en los peores momentos han constituido el principal acicate sobre el que sostenerme. Mi familia siempre estuvo ahí, sin esperar nada a cambio, espoleando mi entereza en aras de la consecución de mis objetivos; al igual que las raíces sostienen al árbol, como la corriente de aire caliente eleva al pájaro cuando vuela, como los remaches cohesionan una estructura consolidada, siempre aportando sus puntos de vista y proveyéndome de la libertad para tomar mis propias decisiones. No podré agradecerles en suficiente medida su infinita paciencia. En especial, dedico este trabajo a la memoria de mi abuelo que, por poco, no ha podido ver el resultado de tanto esfuerzo a pesar de todo lo que luchó.

Tampoco hubiera podido acabar en buen puerto sin la inestimable ayuda y apoyo de Andrés; juntos sufrimos las aventuras y desventuras del *LIFT*. Trabajamos codo con codo para superar las adversidades, vivimos la puesta a punto de un sistema que parecía abocado al fracaso, nos sobrepusimos a las circunstancias y nos divertimos durante el proceso. Matamos muchas células durante estos años, improvisamos medidas desesperadas e ingeniamos métodos para no sucumbir ante la desesperación de los objetivos imposibles. Y más allá, nos conocimos mutuamente, compartimos gustos y aficiones, aprendimos el uno del otro y unificamos puntos de vista para crecer como personas. 'Fue un placer matar células con usted, señor Márquez, aunque ahora ya no se mueren'.

Este trabajo no solo es fruto de los cuatro años académicos en los que se enmarca la tesis. Mi andadura en la ciencia comenzó varios años antes y, con sus más y sus menos, con sus presencias y ausencias, el camino me ha curtido y ha permitido mi evolución en el ámbito personal y profesional. Agradezco enormemente la labor formativa de Pilar, cómo ha sabido y sabe ver mis fortalezas y debilidades; como ha explotado las primeras

y mejorado las segundas; la autonomía, la libertad y la capacidad de decisión; su apoyo durante todo este tiempo. Y no podría olvidarme del resto de personas que forman la 'República Independiente de la Planta Baja del Modular' (incluida la provincia interterritorial del 'secarral'), con la que tantas experiencias he compartido y que me han enseñado buena parte de lo aprendido durante este tiempo. Sin la ayuda y formación que recibí de parte de Laura cuando entré como estudiante, no hubiera podido adquirir las competencias necesarias para llevar a término la tesis. Sin el apoyo de Jenny, y su experiencia con el ARN y sus derivados, no hubiera podido sobrellevar la carga de trabajo en determinados momentos. Begoña Galocha me ha ayudado especialmente con sus críticas constructivas y su buen ojo clínico; sin su ayuda y consejos en la revisión de esta tesis, no hubiera pasado de un panfleto de pensamientos científicos abstractos. Sin las conversaciones distendidas con nuestros vecinos de laboratorio todo hubiera sido soso y gris.

Agradezco a Carlos que aceptara ser mi director en este proyecto conjunto que me permitió optar a la beca de doctorado, y a todo el grupo del Centro Láser, con el que tantos días he compartido. Sin aquellas sobremesas todo hubiera sido muy triste.

Mención especial merece Concha; coincidimos durante un corto periodo, pero su visión de las circunstancias creó un profundo poso en mi manera de pensar. Las conversaciones que mantuvimos en muchos momentos supusieron un claro de cordura, un acicate a la realidad, en medio del confuso caos sin sentido que se cernía y cierne sobre el modular.

Me gustaría agradecer a Sandra Rodríguez y Raúl Torres, de la Unidad de Citogenética Molecular del CNIO, su formación en la utilización del sistema CRISPR/Cas y por compartir sus conocimientos con nosotros. Sin el aprendizaje basado en su experiencia este trabajo no hubiera sido posible.

Por último, agradezco al MECD la financiación recibida, a través de una ayuda FPU, para la realización de esta tesis; no así su trato durante estos cuatro años, ni las 'facilidades' recibidas, ni su soporte a la hora de realizar estancias en el extranjero, ni su percepción de la actividad de los investigadores predoctorales. Vaya desde aquí mi velada crítica para que quede registrada como fotografía de la época que nos ha tocado vivir.

Index

Abbreviations

A

AML1: Acute myeloid leukemia 1 (RUNX1).

AP-1: Activator protein 1.

AS: Ankylosing spondylitis.

ATF-3: Activating transcription factor 3.

B

BFA: Brefeldin A.

Bp: Base pair.

BS(s): Binding site(s).

C

cAMP: Cyclic adenosine monophosphate.

CaRE(s): CRISPR-a responsive element(s).

Cas9: CRISPR associated protein 9, from *Streptococcus pyogenes.*

CCL: Chemokine (C-C motif) ligand.

CCR: C-C motif chemokine receptor.

CD45: Cluster of differentiation 45.

CD69: Cluster of differentiation 69.

CFU: Colony forming units.

Chd4: Chromodomain-helicase DNA-binding protein 4.

ChIP: Chromatin immunoprecipitation.

CNSs: Conserved non-coding sequence(s)/region(s).

Cpf1: CRISPR-associated endonuclease in *Prevotella* and *Francisella* 1.

CREB: cAMP response element-binding protein.

CRISPR: Clustered regularly interspaced short palindromic repeats.

crRNA: CRISPR RNA.

Ct: Cycle threshold.

CTL: Cytotoxic T lymphocyte.

CXCR: C-X-C motif chemokine receptor.

D

DEG: Differentially expressed gene.

DIC: Differential interference contrast.

DSB: Double strand break.

DsiRNA: Dicer-substrate interfering RNA.

E

eGFP: Enhanced green fluorescent protein.

enChIP: Engineered DNA-binding molecule-mediated chromatin immunoprecipitation.

ENCODE: Encyclopedia of DNA elements.

Erg-1/3: ETS related gene 1/3.

ER: Endoplasmic reticulum.

ERAP1: Endoplasmic reticulum aminopeptidase 1.

ERAP2: Endoplasmic reticulum aminopeptidase 2.

Ets: Erythroblast transformation specific.

F

FACS: Fluorescence-activated cell sorting.

FBS: Fetal bovine serum.

FDR: False discovery rate.

FE BA-LIFT: Fluorescence enhanced blister-assisted laser induced forward transfer.

G

GATA: GATA-binding factor.

GAPDH: Glyceraldehyde 3-phosphate dehydrogenase.

gRNA: Guide RNA.

H

HAT(s): Histone acetyltransferase(s).

hCD69: Human cluster of differentiation 69.

HDAC(s): Histone deacetylase(s).

HDR: Homology directed repair.

HLA: Human leukocyte antigen.

HPRT: Hypoxanthine phosphoribosyl transferase 1.

HSPC(s): Hematopoietic stem progenitor cell(s).

I

ICS: Intracellular staining.

IgG: Immunoglobulin G.

IL-2: Interleukin 2.

IMDM: Iscove's modified Dulbecco's medium.

IFN-γ: Gamma interferon.

Io: Ionomycin.

K

KEGG: Kyoto encyclopedia of genes and genomes.

KO: Knockout.

KRAB: Krüppel associated box.

L

LAT1: Large neutral amino acid transporter 1 (CD98).

LCLs: Lymphoblastoid cell lines.

LC-MS/MS: Liquid chromatography with mass spectrometry.

M

mAB: Monoclonal antibody.

mCD69: Mouse cluster of differentiation 69.

MHC: Major histocompatibility complex.

MFI: Mean fluorescence intensity

MOF: Ortholog of drosophila males absent on the first.

MTA3: Metastasis associated 1 family member 3.

mTOR: Mammalian target of rapamycin.

MW: Molecular weight.

MyoD: Myoblast determination protein 1.

N

NCOR1: Nuclear receptor co-repressor 1.

NFkB: Nuclear factor kappa-light-chain-enhancer of activated B cells.

NGS: Next generation sequencing.

NHEJ: Non-homologous end joining.

NK: Natural killer (cell).

NuRD: Nucleosome remodeling deacetylase (complex).

O

Oct1: Octamer transcription factor 1 (POU2F1).

P

P2A: Self cleaving peptide 2A.

PAM: Protospacer adjacent motif.

PCAF: P300/CBP-associated factor.

PCR: Polymerase chain reaction.

PE: Phycoerythrin.

PE/Cy7: Phycoerythrin/Cyanine 7.

PI: Propidium iodide.

PKC: Protein kinase C.

PMA: Phorbol 12-myristate 13-acetate.

POU2F1: POU domain, class 2, transcription factor 1 (Oct1).

Q

qPCR: Quantitative polymerase chain reaction.

R

RCOR1: REST corepressor 1.

RE: Restriction enzyme.

RPMI: Roswell Park Memorial Institute.

RSEM: RNA-Seq by expectation maximization.

RUNX1: Runt-related transcription factor 1 (AML1).

S

$S1P_1$: Sphingosine-1-phosphate receptor 1.

SPI1: PU.1 transcription factor.

SRF: Serum response factor.

ssODN: Single-stranded oligodeoxynucleotide.

STAT: Signal transducer and activator of transcription.

STRING: Search tool for recurring instances of neighbouring genes.

T

TAL1: T-cell acute lymphocytic leukemia 1.

TAP: Transporter associated with antigen presentation.

TCF: T cell factor.

TCR: T cell receptor.

TF: Transcription factor.

TFBS(s): Transcription factor binding site(s).

TFBR(s): Transcription factor binding region(s).

Tip60: Histone acetyltransferase KAT5.

tracrRNA: Trans-activating RNA.

TSS: Transcription start site.

U

UN: Unedited.

UTR: Untranslated region.

V

VP-64: Viral protein 64.

Summary

Resumen

Summary

In this work, the CRISPR/Cas system genome editing toolbox and the fluorescence enhanced BA-LIFT laser bioprinting tools have been used to study the immune system cells behavior. *CD69* is tightly regulated at the transcription level by the CNS2 regulatory region, whose epigenetic marks define as a bivalent regulatory element that match with the previously definition of *CD69* as a bivalent gene. By CRISPR/Cas9 transcriptional assays and genome editing, two major areas within the CNS2 enhancer with antagonistic but complementary regulatory activities *in vivo* were defined in T cells: a Core E region of ~60 bp with a dual repressor and activator function; and a 5′ region adjacent to the Core E (5′C) of ~160 bp where the major activation transcriptional machinery is associated. Deletion of the Core E led to a *CD69* overexpression both at resting and after stimulation, supporting its repressive role at steady state and avoiding *CD69* overexpression upon stimulation. The Core E is enriched in the Oct1 and Chd4 TFs, whose bivalent function have been previously defined. RNA-seq analysis of ΔCore E T cells showed that *CD69* overexpression is associated with an increase of the chemokine receptors CCR1, CCR2 and CCR5 at steady state, and of its ligands CCL3L3, CCL4 and CCL4L1 after stimulation, highlighting the role of CD69 in the regulation of chemokines and its receptors, as has been previously observed in mouse models.

The role of the human *ERAP2* aminopeptidase in the antigen presentation was assessed by generating a functional *ERAP2* KO in B cells. Comparison of the HLA-B*27:05 ligandomes of ERAP$^{+/+}$ and ERAP$^{-/-}$ cells determine that ERAP2 destroy ligands with N-terminal basic residues and peptides with minority anchor motifs at P2 (K and Q). A compensation effect in other peptide positions was observed, since the overall stability of the peptides remained constant.

A newly developed fluorescence enhanced BA-LIFT laser bioprinting technique has been integrated. A laser-based bioprinting device for accurate identification, selection and printing of single or grouped cells within a complex population was developed and validated. FE BA-LIFT has proven to be an efficient and precise tool with high viability and resolution when depositing biological material in hydrogels or scaffolds with different characteristics. An IFN-γ human NK cell reporter has been generated by CRISPR/Cas9 to be used for the validation of the FE BA-LIFT bioprinter.

Resumen

En este trabajo se ha adaptado la herramienta CRISPR/Cas9 y la técnica de bioimpresión laser FE BA-LIFT para el estudio de diferentes aspectos del sistema inmunitario.

La expresión de *CD69* se controla a nivel transcripcional mediante la región reguladora CNS2, cuyas marcas epigenéticas definen como elemento regulador bivalente, concomitante con estudios previos sugiriendo el carácter bivalente de *CD69*. Mediante ensayos de activación transcripcional y deleción con CRISPR/Cas9 se han definido, *in vivo* en células T, dos áreas principales en CNS2 con funciones reguladoras antagónicas pero complementarias: un núcleo transcripcional de ~60 pb con una función represora y activadora dual, y una región adyacente en 5' (5'C) de ~160 pb donde se asocia la principal maquinaria activadora de la transcripción. La deleción del núcleo de CNS2 induce la sobreexpresión de *CD69* tanto en reposo como tras activación, apoyando su papel como represor de *CD69* en estado basal y evitando su sobreexpresión tras la activación celular. El núcleo de CNS2 está enriquecido en Oct1 y Chd4, cuya función bivalente ha sido definida previamente. El análisis de ARN en células T carentes del núcleo de CNS2 muestra que la sobreexpresión de *CD69* se asocia con un incremento en los receptores de quimioquinas CCR1, CCR2 y CCR5 en estado basal, y de sus ligandos CCL3L3, CCL4 y CCL4L1 tras estimulación, subrayando el papel de CD69 en la regulación de quimioquinas y sus receptores, como ha sido descrito en modelos de ratón.

El papel de ERAP2 en la presentación antigénica se ha evaluado generando células B deficientes en ERAP2 funcional. La comparación del ligandoma presentado en las moléculas HLA-B*27:05 de células $ERAP^{+/+}$ y $ERAP2^{-/-}$ muestra que ERAP2 destruye ligandos con residuos básicos en posición N-terminal y péptidos con motivos de unión canónicos en P2 minoritarios (K y Q). Se observa un efecto compensatorio en otras posiciones peptídicas ya que la estabilidad general de los péptidos permanece constante.

Por último, se ha integrado la técnica de bioimpresión FE BA-LIFT desarrollando y validando un dispositivo de bioimpresión para la identificación, selección e impresión de células individuales o en grupos dentro de una población compleja. La herramienta FE BA-LIFT ha demostrado ser eficaz, precisa y resolutiva sin comprometer la viabilidad celular. La bioimpresora se ha validado mediante la generación de células NK humanas reporteras para la citoquina *IFN-γ* por CRISPR/Cas9.

Preface

As in most fields of biological research, understanding of how immune system behaves has developed through a series of stages. First, some scientist discusses a problem that appears irrelevant or esoteric to others. Understanding of the problem progresses slowly until it reaches a point at which an expected observation leads to the realization that the problem can be addressed with the available tools. Understanding of the involved processes increases with the efforts of the research community, leading to clear explanatory models of the measured observations.

At some point, the field reaches a stage at which the models that seemed such a good explanation fall apart, and predictions that were considered obvious are found to be wrong. The only way forward is to reformulate the established paradigms. At this stage, a sense of frustration grows as the problem becomes more and more complex, even though the research effort continues unabated for a while, resulting in thousands of publications which are contradictory or largely descriptive. The more facts we known, the less we understand the processes we are studying.

At this stage, to continue meaningful research, researchers need both to learn how to design effective tools and how to keep their minds clear under adverse circumstances. Working in the knowledge that this will create further confusion is not particularly stimulating, but new developments can interrupt this period of stagnation. Moreover, the human brain can only keep track of a limited number of variables; once the number of components in a system reaches a certain threshold, understanding the biological process without new analytical tools requires geniuses, and they are rare even outside biology. Immunologists, and biologists in general, use several arguments to convince themselves that problems can be solved if one tries hard enough and does another series of experiments; nevertheless, an approach that is inefficient in analyzing a simple system is unlikely to be more useful for a more complex system. And complexity is relative to the degree of understanding.

Finding ourselves in just such a situation, this research aims to adapt the CRISPR/Cas genome editing toolbox for studying several features of the immune system: **1)** To dissect *CD69* transcriptional regulation using both the CRISPR/Cas9 system's genome editing features, and using a modified dCas9 protein to evaluate the regulatory sequences of the

CNS2 enhancer and the regulatory landscape of bound transcription factors and cofactors; **2)** To assess the effect of the loss of the human ERAP2 aminopeptidase in the antigen presentation, by creating a feasible and efficient method to generate an ERAP2 knockout model with CRISPR/Cas9 genome editing; and **3)** To develop a CRISPR/Cas9 method for efficiently generating cytokine fluorescent reporters which combines the power of genome editing with conventional immunological cell selection methods such as FACS, in order to devise new experimental models of immune system cell interactions.

Due to the importance of cell-cell and cell-matrix interactions within the immune system cell network, an additional aim of this work was to develop and validate a laser-based bioprinting device for accurate identification, selection and printing of single or grouped cells within a complex population. Fluorescence and conventional microscopy imaging were implemented in the bioprinting device so that antibody-fluorophore stained cells can be identified, selected and individually printed.

The implementation of this CRISPR/Cas toolbox for studying the immune system opens doors to new methods of study in this field, and seeks to overcome the bottlenecks of conventional analysis techniques. Additionally, the precision and resolution of the laser bioprinting device allows the configuration and generation of interaction models which can help us understand the complex interactions within immune system cell populations.

Aims

A. CRISPR/Cas9 System as a Toolbox to Study the *CD69* Gene Regulation and the Involvement of ERAP2 in the MHC Class I Antigen Presentation

1. Analysis of the role of CNS2 transcriptional enhancer on *CD69* expression, including activation, suppression and editing of CNS2 cis-elements, using CRISPR/Cas9

- New insights on mouse and human *CD69* genomic *locus* and its regulatory elements.
- Characterization of CNS2 enhancement of *CD69* transcription *in silico* and by reporter assays.
- Assessment of the effect of the recruitment of transcriptional activators or repressors to CNS2 sequences on *CD69* protein expression.
- Generation of human T lymphocytes lacking different regions of CNS2.
- Characterization of CNS2 bound transcription factors by ChIP.

2. Analysis of the effect of ERAP2 on the generation of HLA-B*27 ligands

- Human *ERAP2* genomic *locus* characterization.
- Generation of human HLA-B*27:05 B lymphocytes lacking human ERAP2 ($ERAP2^{-/-}$).
- Analysis of $ERAP2^{-/-}$ HLA-B*27:05 B lymphocytes peptidome by mass spectrometry.

B. Laser Bioprinting Techniques as a New Approach to Study the Immune System

1. Bioprinting of immune system cells using FE BA-LIFT

- Properties of sodium alginate and methylcellulose hydrogels during the printing process.
- Assessment of survival, activation and proliferation of FE BA-LIFT printed cells.

2. Generation of fluorescence reporter cells as a tool to analyze the immune interactions upon FE BA-LIFT bioprinting

- Generation of Natural Killer cells activation reporters by inserting the *eGFP* sequence within the *IFN-γ* coding sequence.

Part A:

CRISPR/Cas9 System as a Toolbox for the Study of the Immune Regulation

CRISPR/Cas9 Genome Editing Tools: Present and Future Perspectives

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) is an adaptive immune system of prokaryotes. It allows them to store a record of invading bacteriophages in order to destroy them during subsequent exposures. CRISPR sequences were initially discovered in the *E. coli* genome in 1987 [1], but their function as an immune mechanism against bacteriophages was not elucidated until 2007 [2]. CRISPR associated proteins (Cas) snips foreign DNA into ~30 bp fragments and pastes them into the CRISPR sequences. Another Cas proteins then expresses and processes the CRISPR *loci* to generate CRISPR RNAs (crRNAs). Using sequence homology, this can then guide a Cas nuclease to the specific bacteriophage genetic material located next to the species-specific protospacer adjacent motif (PAM). Then, the CRISPR/Cas9 complex binds to the foreign DNA and cleaves it to destroy the bacteriophage.

The 2015 classification of the CRISPR/Cas system defined 5 types and 16 subtypes, based on the structure of the effector complex that cleaves genomic DNA [3]. The most recent classification (2019) of the CRISPR/Cas system provides an updated evolutionary classification that includes two major classes, 6 types and 33 subtypes [4]. In 2012 it was first demonstrated that the CRISPR/Cas system could be programmed for targeted DNA cleavage *in vitro* [5]; in 2013 it was proven that the system is capable of performing genome editing of cells in a culture [6, 7], and today, less than four years later, PubMed lists more than 2600 CRISPR-related publications detailing work to improve CRISPR/Cas tools specificity, orthogonality, multiplexibility and development of new applications.

The most common CRISPR/Cas system used is the class 2 type II, which consist of the *Streptococcus pyogenes* Cas9 nuclease [8-12] (which recognizes the 5′-NGG PAM sequence). However, the class 2 type V has been described, which includes the Cas12 (Cpf1), as an alternative because it recognizes 5′-TTN PAM sequences, providing improved targeting in AT-rich genomes [13-16]. Two main genome editing applications of the CRISPR/Cas system have been described: **1)** the double-stranded break (DSB) induced by the Cas nuclease can be repaired through the non-homologous end-joining

(NHEJ) mechanism. This is an error-prone process that produces insertions/deletions which disrupt the target gene (**Figure I.1 A**); and **2)** if a repair template with homology arms is supplied together with the CRISPR components, the DSB could be repaired according to this template using the homology directed repair (HDR) mechanism, allowing for precise sequence editing.

Moreover, and in addition to the genome editing properties of the CRISPR/Cas system, its targeting capabilities have been refined to study biological processes without altering the DNA sequence. To do this, researchers have taken advantage of a catalytically dead Cas9 mutant (dCas9) fused to transcriptional activators **[17, 18]** (such as VP-64) or repressors **[19]** (such as KRAB) to activate or deactivate transcription of given genes **[19, 20]**. dCas9 fusion with fluorescent molecules such as eGFP allows the visualization of the genome *loci* of interest, prompting new insights into nuclear organization **[21-24]**. Epitope-tagged dCas9 can also be used to immunoprecipitate and purify a genomic *locus* with its associated proteins or RNA in a process called enChIP **[25-28]**. Moreover, alternative methods have been developed to tag endogenous proteins, thus enabling specific isolation without the need for custom antibodies **[29, 30]**. **Figure I.1** summarizes the genome editing and non-genome editing applications of the CRISPR/Cas system using the wild type Cas9 or the dCas9.

Using different Cas nucleases expands the possible range of PAM sequences and makes it possible to target every *locus* in every genome. It also improves HDR applications, as editing is more efficient when the PAM is close to the desired site of edit. Genome editing with the CRISPR/Cas system is simple and scalable; multiple guide RNAs (gRNAs) targeting different *loci* can be expressed in the same cell or organism **[6, 7]**. This also makes the technology especially suitable for genetic screening. Also, there are now libraries of pooled gRNAs both for gene knockouts and for transcriptional activation/repression assays, allowing the identification of genes that are important to a given phenotype **[31, 32]**. CRISPR/Cas genome editing, together with the resolution of next-generation sequencing (NGS), is the most robust system for genome-wide screening.

Native CRISPR/Cas systems have a variety of enzymes responsible for processing foreign DNA, whereas, for genome editing, only the Cas9 protein (or a suitable variant) is required. gRNA enables the Cas9 to cut a specific genomic *locus* among many possible *loci* by sequence complementarity, but the target must be upstream (5′) of a protospacer adjacent motif (PAM) so that the Csa9 can generate the DSB. In the absence of either the gRNA or the PAM sequence, the Cas9 will neither bind nor cut the target.

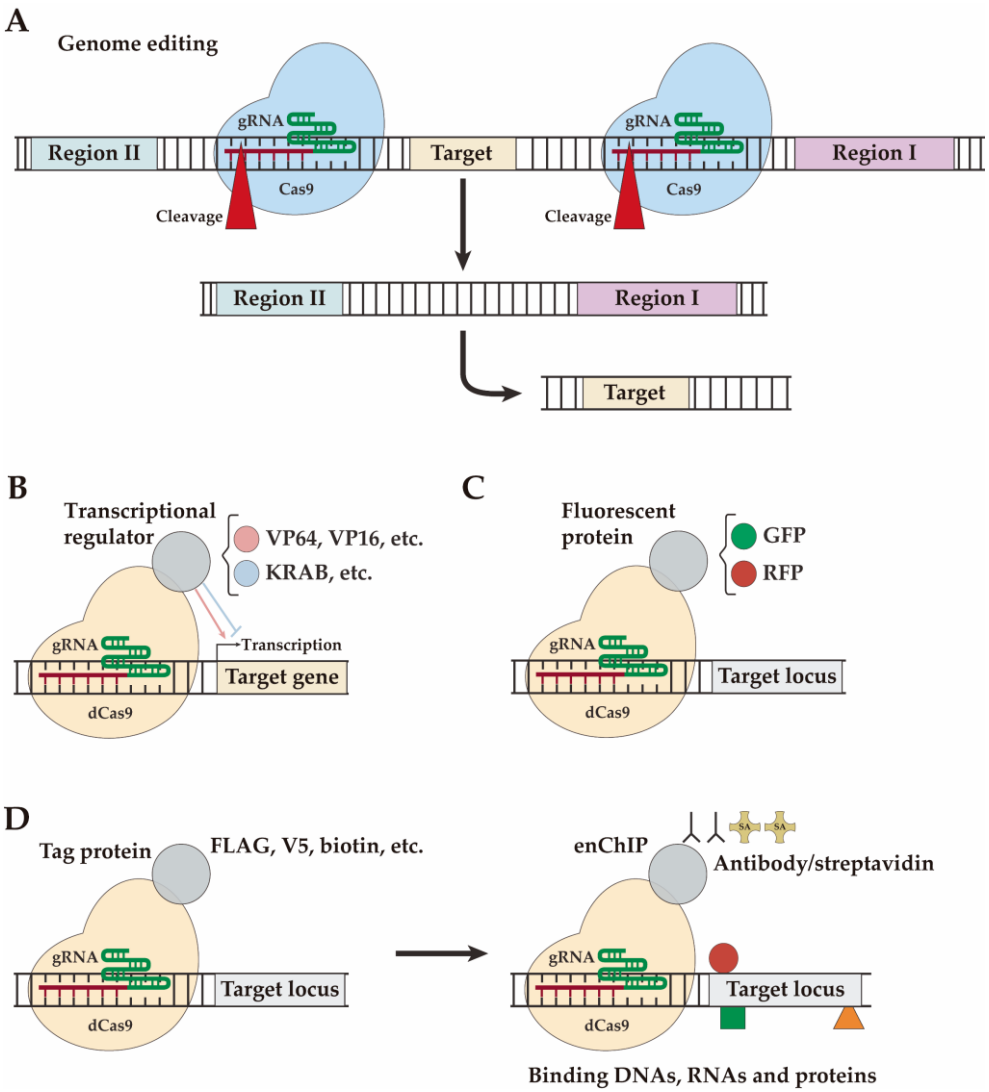


Figure I.1: Genome editing and non genome editing applications of the CRISPR/Cas9 system, using a native Cas9 or a catalytically inactive Cas9 (dCas9), for manipulation and labeling of genome *loci*. A) Recruitment of the Cas9 nuclease to a specific genomic *locus* may result in a double strand break; the resulting cleavage may release the target sequence, so that the uncut extremes can come together**. B)** The dCas9 protein fused to a transcriptional activator/repressor can function as a transcriptional regulator**. C)** The dCas9 protein fused to fluorescent proteins may allow the imaging of specific genomic *loci* in living cells**. D)** The dCas9 protein can be tagged, so anti tag antibodies can be further used to capture targeted genome sequences with its associated DNAs, RNAs and proteins (transcription factors and cofactors) in a method called enChIP. Adapted from [33].

Cas9 protein has two endonuclease domains: The N-terminal RuvC-like nuclease domain and the HNH-like nuclease domain near the protein core. Upon binding, a conformational change positions the nuclease domains to cleave opposite strands of the target DNA, thus resulting in a DSB generation within the ~3-4 nucleotides upstream of the PAM sequence. The native class 2 type II CRISPR/Cas9 system recruits the nuclease to its target with the aid of two RNAs: the crRNA that defines the genomic target, and the transactivating RNA (tracrRNA) that acts as a scaffold linking the crRNA and the Cas9, and also facilitates the processing of mature crRNAs from pre-crRNAs derived from CRISPR arrays. To increase the system performance, researchers condensed these two RNAs into one RNA sequence called single guide RNA (sgRNA) containing both the 20-nucleotide target sequence and the scaffolding sequence necessary for Cas9 binding. Moreover, recent advances have developed the *in vitro* combination of the sgRNA with the Cas9, thus generating a ribonucleoprotein (RNP) that can be easily driven into the cell nucleus to perform the genome editing.

In this work, we have adapted the CRISPR/Cas9 toolbox to study several aspects of the immune system. A mix of the CRISPR/Cas9 tools were used to study the human *CD69* gene *locus*: **1)** the regulatory region CNS2, as well as its main regulatory sequences, were deleted *in vivo* in Jurkat T cells; and **2)** we used the dCas9 protein fused to a transcriptional activator or repressor to assess the contribution of every CNS2 region to the regulatory landscape of *CD69* transcription. The CRISPR/Cas9 genome editing capability was used to generate a functional ERAP2 knockout cell line by deleting exons 5 and 6. Finally, the HDR pathway after DSB induced by the Cas9 was employed, together with two different HDR donors, to generate a Natural Killer (NK) cytokine reporter cell, which express the *eGFP* fluorescent protein together with the *IFN-γ*.

Chapter I:

Core Activity of the CNS2 Enhancer Acts as a Switch to Repress or Activate Human *CD69* Expression

Introduction I

1. *CD69* Physiology and Function

CD69 is a type-II membrane glycoprotein with a C-type lectin-like domain [34] which has been described as an early leukocyte activation marker. CD69 plays an important role in the regulation of immune responses, as its blockade increases mTOR signaling pathway that leads to cytokine and chemokine production, cell proliferation, increased survival and massive leukocyte mobilization [35-39]. CD69 is dully expressed on the membrane of some leukocyte subsets in a steady state [40-42], and it is strongly and rapidly upregulated on all leukocytes upon activation. Both *in vitro* and *in vivo*, *CD69* is upregulated in response to TCR engagement [43, 44], cytokines [45, 46] and PKC activators [41, 47-50]. This upregulation is strongly dependent on transcriptional induction [43, 49, 51].

By studying of $CD69^{-/-}$ mice, *CD69* has been defined as a modulator of the immune response by controlling cytokine and chemokine production and leukocyte circulation in different murine models of tumor, infection, autoimmune disease and other inflammatory models [36, 38, 52-56]. Although the *CD69* expression pattern suggest its pro-inflammatory effect, recent reports indicate that CD69 exerts a function aimed to limit the inflammatory phenomenon and tissue damage. $CD69^{-/-}$ mice exhibit few immune abnormalities in steady state conditions, but they are prone to developing an increased anti-tumor and infection response, and several inflammatory diseases, mainly those mediated by Th17 lymphocytes [38, 53, 57, 58].

Figure I.2 summarizes the structure of CD69 protein, as well as its described ligands and cis interactions. Galectin-1, oxidized LDL (oxLDL), S100A8/A9 (calprotectin), and myosin light chains 9 and 12 (Myl9/12) have been proposed as CD69 ligands [39, 58, 59] in T cells. It has been shown that the carbohydrate-dependent interaction with galectin-1 is able to inhibit Th17 differentiation in a Th17 polarization culture of $CD4^+$ T cells, while CD69-deficient cells were unresponsive [60], and oxLDL interactions with CD69 diminished the percentage of $IL\text{-}17^+$ and $IFN\text{-}\gamma^+$ cells generated in response to Th17- or Th1-polarizing stimuli, thus favoring Treg differentiation [61].

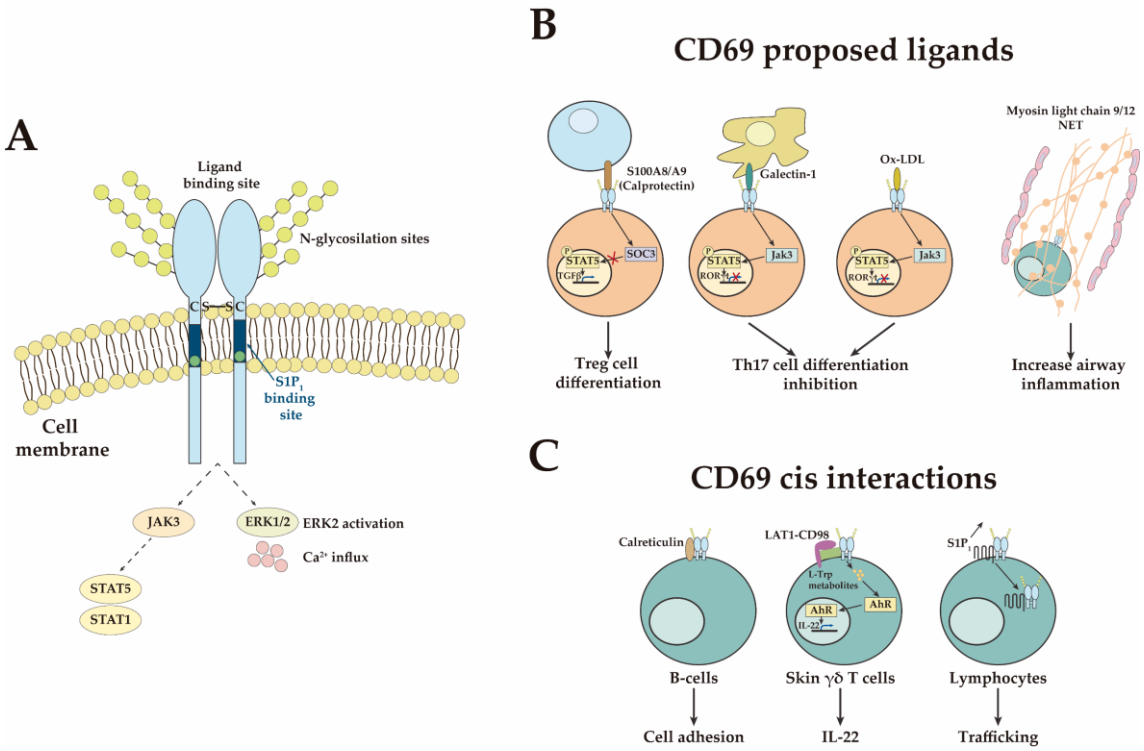


Figure I.2: CD69 structure, signaling pathways and molecular interactions. A) CD69 molecule structure and signaling pathways. **B)** CD69 proposed ligands. Interaction with S100A8/A9 (calprotectin) has been described to promote Treg cell differentiation. Interaction with galectin-1 and oxidized LDL (oxLDL) leads to a Th17 cell differentiation inhibition, favoring the Treg cell differentiation. Interaction with Myl9/12 regulates airway inflammation by recruiting activated T cells to the lungs. **C)** CD69 described cis-interactions. CD69-calreticulin interaction has been described to have implications in cell adhesion, especially in B cells. CD69 interaction with L-tryptophan transporter LAT1-CD98 regulates the L-tryptophan dependent activation of the AhR, leading to IL-22 production in γδ T cells. CD69 regulates thymocyte and lymphocyte egress from lymphoid tissues through $S1P_1$ interaction and internalization, inhibiting lymphocyte egress from the thymus and peripheral lymphoid organs.

S100A8/A9-CD69 (calprotectin-CD69) axis has been described as indispensable for the Treg cell differentiation, an interaction that is carbohydrate dependent [62]. CD69 interaction with Myl9/12 recruits activated T cells into the inflammatory lung, positively regulating airway inflammation [34, 63]. CD69 cis-interaction with several lymphocyte surface proteins has been described. Calreticulin at the cell surface of activated PBMCs has been found associated with CD69, mostly in B cells, and may have important implications in cell adhesion [64]. CD69 cis-interaction with the aromatic amino acid transporter complex LAT1-CD98 has been shown to regulate L-tryptophan uptake and intracellular aryl hydrocarbon receptor (AhR)-dependent IL-22 secretion in γδ T cells [65].

CD69 regulates thymocyte and lymphocyte egress from lymphoid tissues through its interaction in cis with the sphingosine-1-phosphate receptor 1 ($S1P_1$). CD69 binds to $S1P_1$ through its intracellular domain (**Figure I.2 A**), which results in the internalization of the CD69-$S1P_1$ complex **[66, 67]**, thus inhibiting lymphocyte egress from the thymus and peripheral lymphoid organs **[56]**. Also, through $S1P_1$, CD69 blockade induces rapid mobilization and high proliferation of hematopoietic stem and progenitor cells (HSPCs) **[35]**. In the absence of infection, human CD69 targeting with α-CD69 mAb induces an immune activation characterized by the mobilization, proliferation, enhanced survival, and marked production of several proinflammatory cytokines by immune cells **[54]**. Moreover, α-CD69 mAb treatment induces the mTOR signaling-dependent expansion and egress of HSPCs from the bone marrow **[35]**, and CD69 targeting has been proposed to enhance vaccine protectiveness **[68]**.

2. *CD69* Gene Regulation

In mammals, 95% of the genome is noncoding, and 40% of noncoding regions plays a role in transcriptional regulation **[69]**. In combination with gene promoters, cis-regulatory elements contribute to gene transcriptional regulation as enhancers or repressors. Cis-regulatory elements are often identified as noncoding sequences with DNase I hypersensitivity; they also contain binding sites for trans-acting transcription factors (TFBSs) and serve as centers of epigenetic changes **[70, 71]**. These elements can be located around or within genes and can affect their transcription from as far as megabases **[72]**.

Enhancers are acetylated DNA elements of a few hundred base pairs long that interact with transcription factors (TFs) and promoters. Enhancers can exert their function in an orientation and distance-independent manner, and are thus capable of targeting both upstream and downstream genes **[73]**. It is generally accepted that enhancers can be found at different epigenetic states **[74]**: **i)** active enhancers are decorated with histone H3 lysine 27 acetylation (H3K27ac) and histone H3 lysine 4 monomethylation (H3K4me1) epigenetic marks**, ii)** primed enhancers are not enriched for H3K27ac but retain H3K4me1, driving basal levels of gene activation**,** and **iii)** poised enhancers are

marked with both canonical marks for active enhancers (H3K4me1 and H3K27ac) and with the repressive histone H3 lysine 27 trimethylation (H3K27me3) mark [75-77].
Moreover, some enhancers behave as bivalent domains that are distinguished as having two histone marks that are associated with both positive (H3K4me3 [73]) and negative (H3K27me3) transcriptional events, a characteristic that imposed a more nuanced poised state on the chromatin instead of simply 'on' or 'off' [78]. These regulatory regions determine cell identity in both healthy and pathological states. For example, some enhancers were identified as differentially activated in certain tumor cells [79]. Also, it is suspected that enhancers of immune system genes may be involved in the regulation of inhibitory checkpoint expression, thus determining the physiology and pathology of the immune system cell subsets. The three-dimensional conformation of the genome seems to be important to promote contacts between distant regulatory elements and their target genes. Thus, these sequences form complex regulatory landscapes within the noncoding genome.

CD69 gene is encoded in the NK complex, located in human chromosome 12 and mouse chromosome 6 [51, 80]. Human and mouse *CD69* promoters are able to direct transcription in resting cells, as well as to increase *CD69* transcription in stimulated cells [81]. Induction of *CD69* transcription depends on its promoter, which binds cis-regulatory elements such as Erg-1, Erg-3, ATF-3/CREB, AP-1 and NFkB TFs [81-83]. In addition to *CD69* promoter, the existence of another five cis-regulatory regions has been previously described: A non-conserved hypersensitivity site (HS) located within the first intron of the *CD69* gene and four conserved noncoding sequences (CNS1-4) upstream of the *CD69* promoter [84, 85]. Murine CNS1-4 are responsible for developmental-stage and lineage-specific regulation of *CD69* expression in T cells, but not in B cells. These regions showed DNaseI sensitivity and specific histone modifications; moreover, CNS2 and CNS4 had constitutive and inducible enhancer activity in reporter assays, being CNS2 the region with the highest inducible enhancer capacity [84, 85].
Recent work proposed CNS2 as the main enhancer of *CD69* promoter transcription *in vivo* [86] based on the characterization of TFBSs conserved among six mammal species

and explored the contribution of some of its regulatory features to its enhancing capacity. Furthermore, recruitment of a dCas9-VP64 with gRNAs targeting all the PAMs located 100kb upstream CD69 TSS supports the role of CNS2 as a stimulation-responsive immune enhancer **[18]**. Nevertheless, those experiments did not take into account either the gRNA efficiency or the sequence that exerts the *CD69* transcriptional activation.

This work presents a comprehensive map of the TFs that bind to mouse and human CNS2. A 60 bp sequence within CNS2 is proposed as the enhancer Core (Core E), defined by the recruitment of the dCas9-VP64 transcriptional activator, whereas a high repression is induced when recruiting the dCas9-KRAB transcriptional repressor to the 5' adjacent region (region 5'C). CRISPR/Cas9 deletion of the Core E increases *CD69* expression 2.5 times, highlighting a dual role of this regulatory sequence on *CD69* transcriptional regulation. A deeper characterization of the TFs that exerts this dual role was carried out by ChIP analysis. Results suggest that *CD69* expression is controlled by CNS2 through the Core E region with a dual repressor and activator function and that in the 5'C region the major transcriptional machinery is associated, with an activation function.

Materials and Methods I

1. Characterization of Both Human and Mouse *CD69*

Sequences and genomic *locus* distribution of human and mouse *CD69* were obtained from the UCSC Genome Browser [87] (ENCODE consortium [88]). Mouse (GRCm38/mm10 assembly, Chr6: 129,259,289 - 129,321,308) and human (GRCh37/hg19 assembly, Chr12: 9,900,752 - 9,952,771) *CD69* genomic *locus* were analyzed with the VISTA browser [89-91] to identify CNSs and similarities. Transcription factor binding data, determined by ChIP-seq in different cell lines, were obtained from the ENCODE consortium, the ChIPBase v2.0 [92] (version 2.3.4, Qu Lab, School of Science, Sun Yat-Set University, Guangzhou, China) and the ReMap [93, 94] databases for human data; and with both the ChIP Atlas [95] (Database Center for Life Science, Kyushu University, Fukuoka, Japan) and the ChIPBase v2.0 databases for mouse data. TFBSs were displayed using Snapgene® to create a comprehensive map of the regulatory features of *CD69*.

2. Luciferase Plasmids

The mouse *CD69* promoter (-1 to -609, BAC clone RP24-188C4) was cloned into *Bgl*II and *Hind*III restriction enzyme (RE) cloning sites of the commercial luciferase vector pGL3 basic (Promega, E1751, Madison, WI, USA, [**sequence**]) (pPr plasmid). Then, the mouse CNS2 region (mouse genome assembly MGSCv37 (mm9) chr6:129,234,359-129,235,318) was cloned into *Kpn*I and *Xho*I RE sites, with the introduction of an *Eco*RI site by *Kpn*I for further cloning (pPr2 plasmid). Modified CNS2 constructs containing different regions or specific TFBSs were PCR amplified from the pPr2 plasmid employing custom primers (**Table I.1**), followed by cloning of the fragments into *Eco*RI and *Xho*I sites in the plasmid containing the mouse *CD69* promoter (**Table I.2**). The resulting luciferase plasmids (**Figure SI.1**) were validated by Sanger sequencing.

3. Luciferase Assays

5-7×10^5 human Jurkat T cells were transfected with 1 µg of modified firefly luciferase plasmid (purified with the Plasmid Maxi Kit, QIAGEN 12163, following manufacturer's protocol) plus 20ng of pRL-TK (Renilla luciferase plasmid, Promega E224, [**sequence**]) to standardize the luciferase activity independently of the transfection efficiency

between samples, using Effectene (QIAGEN 301425) following manufacturer's protocol. After transfection, cells were cultured at 37°C and 5% CO_2 for 24 hours and then stimulated with 10 ng/ml PMA plus 500 ng/ml ionomycin for an additional 24 hours. Forty-eight hours after transfection, cells were lysed with passive lysis buffer (Promega E1941), and luciferase activity (firefly/renilla) was measured with a Dual-Luciferase Kit (Promega E1910) using the Orion II microplate luminometer (Berthold 11300010, Bad Wildbad, Germany).

4. Dicer Substrate Small Interfering RNA (DsiRNA) Design, Transfection, and CD69 Characterization

All DsiRNA described were synthesized by IDT (Coralville, IA, USA); three DsiRNA were designed for each gene to be silenced, in addition to two additional DsiRNAs and a negative control targeting NC1. The reference genome used for DsiRNA design was the mouse assembly GRCm38 (mm10). **Table I.3** provides the sequence and cross-reacting properties of DsiRNAs against *Oct1* and *RUNX1* mRNAs.

For mRNA interference experiments, 10^5 mouse EL-4 T cells were transfected either with a 1 μM pool of three DsiRNAs against *Oct1* or *RUNX1* in a 1:1:1 ratio, or with a negative control DsiRNA against NC1. RNA was extracted (**M&M I.11**) and a qPCR (**M&M I.13**) was performed to asses mRNA levels of *RUNX1* and *Oct1*.

Transfected cells were cultured in a 96 well plate at 37°C and 5% CO_2 for 16 hours and then stimulated, or not, with 10 ng/ml PMA and 500 ng/ml ionomycin for an additional 6 hours. Twenty-two hours after DsiRNA transfection, *CD69* surface expression was measured by flow cytometry, as described in **M&M I.16**.

The percentage of *CD69* expression inhibition with respect to its expression in the NC1 control was represented using GraphPad Prism 7 (version 7, San Diego, CA, USA).

5. crRNA Design and CRISPR/Cas9 Ribonucleoprotein (RNP) Transfection

crRNA design

crRNA analysis was carried out using Benchling (Biology software, 2019. Retrieved from benchling.com). Human CNS2 sequence (chr12: 9,764,620 - 9,765,633, GRCh38/hg38 assembly,) was analyzed using the CRISPR/Cas9 design tool from Benchling.

Only crRNAs with an off-target score [96] higher than 70 were considered, and crRNAs with the highest on-target score [97] were selected. Table I.4 shows the designed crRNAs, their sequences, targets and genomic locations.

Selected crRNAs were tested with the OligoAnalyzer Tool (Integrated DNA Technologies) so that they had 40-60% GC, a lower probability to form hairpin at working temperature, and a lower self-dimer probability.

sgRNA generation

crRNA and tracrRNA were synthesized by IDT (Alt-R CRISPR-Cas9 crRNA and Alt-R CRISPR-Cas9 tracrRNA, 1072533) and resuspended in nuclease-free duplex buffer (IDT, 11-05-01-12) to a final concentration of 100 µM. Cas9-V3 nuclease (61 µM) was purchased from IDT (Alt-R® S.p. Cas9 Nuclease V3, 1081059) and diluted with an equal volume of nuclease-free duplex buffer.

For sgRNA generation, 1.5 µl of 100 µM crRNA plus 1.5 µl of 100 µM tracrRNA and 1 µl of nuclease-free duplex buffer were mixed and heated at 80°C for 10 minutes in the thermocycler. sgRNAs were tempered at room temperature before RNP assembly.

For RNP complex formation, sgRNA and Cas9-V3 were mixed in a 1:1 ratio by gentle stirring and incubated at room temperature for 20 minutes.

CRISPR/Cas9 RNP transfection

For the deletion of CNS2 regions, 2×10^5 Jurkat T cells were transfected with 1 µl of RNP complex using the Neon Transfection System (ThermoFisher Scientific, MPK5000 and MPK1096) programmed as detailed in Table I.5.

For transcriptional activation and repression assays, 2×10^5 Jurkat CRISPR-a or CRISPR-i, respectively, were transfected with 1 µl of sgRNA targeting different sequences of CNS2 (M&M I.7).

6. CNS2 Transcription Factor Binding Sites (TFBSs) Deletion

For deletion of the whole CNS2 region and the sequences containing the main TFBSs, 2×10^5 Jurkat T cells were transfected with the CRISPR/Cas9 RNP containing the crRNAs described in M&M I.5 in the combinations depicted in Table I.6.

Forty-eight hours after transfection, DNA was extracted and the CNS2 *locus* was analyzed by PCR using primers that amplified outside and inside the deleted region (**Table I.7**). Transfected cells were then cloned, CNS2 expression was analyzed by PCR (**Figure SI.3**) and amplicons were sequenced, allowing the selection of homozygous clones for the selected deletion.

Deletion of CNS2 and the sequence containing Oct1 BSs in Jurkat CRISPR-a was carried out as explained above but, before the cloning strategy, cells were transfected with the sgRNA CNS2_Oct1_A (6) and 24 hours later those $CD69^-$ were selected and sorted by FACS (**M&M I.16**). Then, cells were cloned and CNS2 *locus* was PCR amplified in order to find homozygous clones for each deletion. CD69 protein expression in Jurkat and Jurkat CRISPR-a cell clones lacking the whole CNS2 sequence or sequences containing the main TFBSs were analyzed by flow cytometry (**M&M I.16**).

7. Transcriptional Activation and Repression in T Cells by Using Modified Cas9 Nucleases

Two different Jurkat T cell lines were used for transcriptional activation and repression assays, kindly provided by Dr. Alexander Marson (University of California, San Francisco, CA, USA). For transcriptional activation assays, 2×10^5 Jurkat CRISPR-a T cells which constitutively express the dCas9 protein fused to a transcriptional activator and the eGFP (dCas9-VP64-eGFP) were transfected with gRNAs targeting different sequences within CNS2 enhancer (**Table I.4**, crRNAs 2-16), with a gRNA targeting a sequence outside but near CNS2 (**Table I.4**, crRNA 1), or a control gRNA targeting the *HPRT* gene. CD69 protein expression was measured by flow cytometry (**M&M I.16**) 3, 6, and 24 hours after gRNA transfection. For transcriptional repression assays, 2×10^5 Jurkat CRISPR-i T cells, which constitutively express the dCas9 protein fused to a transcriptional repressor (dCas9-KRAB) were transfected with the gRNAs described above. Eighteen hours after transfection Jurkat CRISPR-i cells were stimulated with 50 ng/ml PMA and 1 μg/ml ionomycin for an additional 6 hours. Twenty-four hours after transfection CD69 protein expression was measured by flow cytometry (**M&M I.16**).

8. Chromatin Immunoprecipitation (ChIP)

ChIP was performed on resting Jurkat T cells or 24 hours after stimulation with 50ng PMA and 1 μg ionomycin. Cells were crosslinked with 1% formaldehyde for 10 minutes and then lysed. Chromatin was sheared by sonication using a Bioruptor® Pico sonication device (Diagenode, B0160010). Sheared chromatin was immunoprecipitated with a 1:1 equimolecular mix of antibodies α-Chd4 (Abcam, ab70469 and ab72418), a 1:1 mix of α-Oct1 (Bethyl Laboratories, A-301-716A and A-301-717A), α-RUNX1/AML-1 (Abcam, ab23980), α-SPI1 (Cell Signaling 2266S), or control α-IgG (Cell Signaling 2729). Precipitated DNA was isolated with the QIAquick PCR purification kit (QIAGEN, 28106) and analyzed by qPCR for enrichment on CNS2 sequence using custom PrimeTime® qPCR probes (Integrated DNA Technologies) as shown in **Table I.8**.

9. Genomic DNA Extractions

To perform DNA extractions, cells were centrifuged at 6,000×g for 5 minutes, and the cell pellet was dissolved in 200 μl of DNA extraction buffer (0.1M Tris-HCl pH 7.5, 0.05M EDTA pH 8.0, 1.25% SDS) with 0.2 μg/μl RNAse A. DNA extract was incubated at 37°C for 30 minutes and then proteinase K at 0.2 μg/μl was added, followed by incubation at 55°C for 1 hour. DNA of digested cells were precipitated as follows: 400 μl of absolute ethanol (EtOH) and 2 μl sodium acetate 3M pH5.2 (NaAc 3M) were added to cell lysate, followed by centrifugation at maximum speed for 10 minutes. The supernatant was discarded, DNA pellet was washed with 70% EtOH and centrifuged at maximum speed for an additional 2 minutes. The supernatant was discarded and the DNA pellet dried until ethanol evaporated. The DNA was dissolved in 20 μl of H_2O, and its concentration and purity were analyzed with a Nanodrop ND-1000 (ThermoFisher) spectrophotometer.

10. Genomic Genotyping PCR

The NZYTaq 2x Colorless Master Mix (NZYTech, MB35703) was used accordingly to the manufacturer's instructions. Briefly, in a 0.2 ml Eppendorf tube were mixed 5 μl of 2x Master Mix, 1 μl of primer mix (each) and 1 μl of DNA (10-50 ng) in a final reaction volume of 10 μl.

The PCR products were resolved in agarose gels (1-4%) containing 1 μL of GreenSafe Premium (NZYTech, MB13201) per 50 ml agarose gel and visualized with UV light.

11. Total RNA Extractions

Total RNA was isolated using the NZY Total RNA Isolation Kit (NZYTech, MB13402) following the manufacturer's instructions. Briefly, a minimum of 10^6 cells were pelleted at 6,000×g for 5 minutes at 4°C, and 350 μl of buffer NR and 3.5 μl of β-mercaptoethanol were added. The cell pellet was vortexed vigorously, transferred to a NZYSpin homogenization column and centrifuged for 1 minute at 11,000×g. 350 μl of 70% EtOH was added to the flow through and mixed immediately by pipetting up and down. The lysate was then loaded in a NZYSpin binding column and centrifuged at 11,000×g for 3 minutes. Flow through was discarded, 350 μl of buffer NI was added to the center of the column and centrifuged again at 11,000×g for 30 seconds. Each isolation was then incubated at room temperature for 15 minutes after adding 95 μl of digestion mix (90 μl digestion buffer plus 10 μl DNase I). 200 μl of buffer NRW1 was added to the column and it was centrifuged for 1 minute at 11,000×g. Then two washes with 600 μl and 250 μl of buffer NRW2 were performed, and the column was transferred to a new 1.5 ml Eppendorf tube. 10-20 μl of RNase-free water were added, incubated for 1 minute at room temperature and centrifuged at 11,000×g for 1 minute to elute the RNA.

The extracted RNA concentration and purity was analyzed with a Nanodrop ND-1000 (ThermoFisher) spectrophotometer.

12. RT-PCR

DNase I digestion

DNase I digestion of isolated RNA was performed before RT-PCR using the RQ1 RNase-Free DNase (Promega, M6101) kit following the manufacturer's instructions. Briefly, 1-5 μg of RNA was mixed with 2 μl of RQ1 RNase-Free DNase 10× reaction buffer and 1 u/μg RNA of RQ1 RNase-Free DNase in a final volume of 20 μl. Samples were incubated at 37°C for 30 minutes and then 1 μl of RQ1 DNase stop solution was added to end up the reaction. Samples were incubated at 65°C for 10 minutes to inactivate the DNase.

RT-PCR

To obtain cDNA, 500-1000 ng of DNase treated RNA was used. The reverse transcriptase reaction was performed using the NZY First Strand cDNA Synthesis Kit (NZYTech, MB12502) following the manufacturer's instructions. RNA was mixed with 10 µl of NZYRT 2x Master Mix and 2 µl of NZYRT Enzyme Mix in a final volume of 20 µl. The reaction was gently mixed and incubated in a thermocycler with the following program:

1. 25°C, 10 minutes.
2. 50°C, 30 minutes.
3. Inactivation at 80°C for 5 minutes.
4. Chill at 4°C.

Then 1 µl of NZY RNase H (*E. coli*) was added to the reaction, and it was incubated at 37°C for 20 minutes. cDNA was used directly in qPCR.

13. Quantitative PCR (qPCR)

TaqMan qPCR

TaqMan probes were purchased from IDT as PrimeTime® qPCR assay. 50 ng of cDNA was mixed with 0.5 µl of PrimeTime® qPCR assay 20× and 5 µl of NZY qPCR Probe (2x) ROX plus (NZY Technologies, MB225) in a final volume of 10 µl. Negative controls and samples were run in the QuantStudio 3 Real Time PCR system (Thermo Fisher Scientific) using the program described in **Table I.9**.

To calculate the relative expression of different genes, the $2^{-\Delta\Delta Ct}$ method was used. Briefly, the Cts of the gene of interest (target) and the Cts of the reference gene (human *β2microglobulin* or mouse *GAPDH*) were compared ($\Delta Ct = Ct_{target} - Ct_{reference}$). Then, the ΔCt of each sample was normalized to the control sample ($\Delta\Delta Ct = \Delta Ct_{target} - \Delta Ct_{control}$).

SYBR green qPCR

For SYBR green qPCR, 50 ng of cDNA was mixed with a 2.5 µM mix of primers and 5 µl of NZY qPCR Green Master Mix (2x) ROX plus (NZY Technologies, MB222) in a final volume of 10 µl. Negative controls and samples were run in the QuantStudio 3 Real Time PCR system (Thermo Fisher Scientific) using the program described in **Table I.10**. A melting curve stage was introduced to analyze the presence of primer dimers.

14. RNAseq Analysis

For gene expression analysis, the clean reads were mapped to the reference using Bowtie2 v2.2.5 **[98]**, and then gene expression level was calculated with RSEM **[99]**, a package for estimating gene and isoform expression levels from RNA-seq data. DEGs were detected with a Poison distribution as previously described **[100]**. With the KEGG annotation results, DEGs were classified and the pathway functional enrichment was done using R's `phyper` function. FDR for each p value was calculated, and FDR values not larger than 0.01 were defined as significantly enriched. DEGs were mapped to the STRING **[101]** database to obtain a network interaction diagram of upregulated gene expression (FDR < 10^{-3}). DEGs retrieved by the STRING analysis were validated by qPCR using the primers depicted in **Table I.11.**

15. Sequencing

Plasmid constructions and PCR products were purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey Nagel, 740609.240C) and quantified by NanoDrop ND-1000. DNA was mixed with primers and sequenced in the Genomics Unit of the Spanish National Center of Microbiology. Sequences were analyzed with Snapgene (from GSL Biotech; available at snapgene.com) and the BlastN tool **[102, 103]**.

16. Flow Cytometry Analysis and Cell Sorting

CD69 staining was performed as follows: 10^5 cells were washed twice in cold 1X PBS and stained for 30 minutes at 4°C in the dark with 0.5 µg of an α-mouse CD69/PE-Cy7 monoclonal antibody (Clone H1.2F3, eBioscience 25-0691-81), or 2 µl of an α-human CD69/PE monoclonal antibody (Clone FN50, Biolegend, 310906). CNS2 genome-edited clones were stained with an α-hCD69/PE monoclonal antibody, as stated above and sorted to select $CD69^{-/-}$ or $CD69^{+/+}$ cells.

Flow cytometry samples were acquired with a FACSCanto (Becton Dickinson, Franklin Lakes, NJ, USA) flow cytometer, whereas FACS samples were acquire with a FACSAria flow cytometer. Data were analyzed and represented using FlowJo™ software.

17. Statistical Analysis

GraphPad Prism version 7.0 (GraphPad Software Inc., San Diego, CA, USA) was used to perform one-way or two-ways analysis of variance (ANOVA), or t-tests, where indicated. All data are presented as the mean ± SD. The criterion for significance was set as described for each test performed.

18. Tables

Table I.1: List of primers used to amplify the CNS2 sequences for further cloning and sequencing. Sequences in italics indicate the extra bases that allow restriction enzymes to cut; bold sequences indicate *Xho*I and *Eco*RI target sequences; plain text indicates the CNS2 sequence.

Primer	Sequence 5′ → 3′
I_Fw	*GAG***GAATTC**TTTCAGAAACCTCCCTCCGG
I_Rv	*AGA***CTCGAG**CAGCATCTGTTGTGATTAGCAT
I+II+III_Fw	*CC***GAATTC**CCAGTACCTCATTCACTCACTG
I+OCT_Fw	*AAAA***GAATTC**AGCTGCCGTGATAAGGACTGTA
III_Fw	*ATAAA***GAATTC**CCAGTACCTCATTCACTCACTG
I+OCT_Rv	*AGT***CTCGAG**CACTAAAACCCAGTGCAGTTCT
III+OCT_Fw	*AAA***GAATTC**CCAGTACCTCATTCACTCACTG
OCT_Fw	*AAAA***GAATTC**AGCTGCCGTGATAAGGACTGTA
OCT_Rv	*AAA***CTCGAG**CTGAAAGGAAATGATGTAATGT
III_Rv	*AAA***CTCGAG**GCCTTAGTGGTATGTGTTCAAAG
I+II_Fw	*AAAA***GAATTC**TGGTTCCTTTGAACACATAC
III-GATA_Rv	*AAAA***CTCGAG**AGGAAGGCAGGGTGT
MyoD+GATA_Rv	*AAA***CTCGAG**AGGAAGGCAGGGTGTACA
RV3_Fw *	CTAGCAAAATAGGCTGTCCC
Prom_Seq_Rv *	TGACATGGGAAAAGCACTGGA

* indicates the primers used for sequencing and validation of the resulting plasmids.

Table I.2: Primers used for amplification of luciferase plasmids and length of the expected amplicon.

Construct	Primers used	Amplicon length
I+II+III	I+II+III_Fw + I + Oct1_Rv	297 bp
II+III	III+Oct1_Fw + Oct1_Rv	227 bp
I+II	Oct1_Fw + I + Oct1_Rv	177 bp
I	I_Fw + I+Oct1_Rv	70 bp
II	I+II_Fw + Oct1_Rv	107 bp
III	III_Fw + III_Rv	120 bp
GATA+MyoD	I+II_Fw + MyoD + GATA_Rv	65 bp
Oct1	Oct1_Fw + Oct1_Rv	60 bp
GATA+MyoD+III	III_Fw + III-GATA_Rv	156 bp
I+Oct1	I+Oct1_Fw + I+Oct1_Rv	142 bp

Table I.3: Nucleotide sequences of the sense and antisense strands and cross-reaction of the DsiRNAs used to silence *RUNX1* and *Oct1 mRNAs*.

DsiRNA	Sequences	Cross-reacting transcript
Runx1.13.1	**5′-** AAGAAAGAUAUCAAGUACUACAUtt **-3′** **3′-** AAUUCUUUCUAUAGUUCAUGAUGUAAA **-5′**	NM_001111022 NM_001111021 NM_001111023 NM_009821
Runx1.13.2	**5′-** AUGGCAGGCAACGAUGAAAACUAct **-3′** **3′-** AGUACCGUCCGUUGCUACUUUUGAUGA **-5′**	NM_009821 NM_001111023 NM_001111021 NM_001111022
Runx1.13.3	**5′-** GAAGAACCAGGUAGCGAGAUUCAac **-3′** **3′-** UACUUCUUGGUCCAUCGCUCUAAGUUG **-5′**	NM_001111022 NM_001111021 NM_001111023 NM_009821
Oct1.13.1	**5′-** UAAAUUUCAUGAAAGCUUUACUUgt **-3′** **3′-** CGAUUUAAAGUACUUUCGAAAUGAACA **-5′**	NM_198932 NM_011137 NM_198934 NM_198933
Oct1.13.2	**5′-** CAAGAAUGAAUAAUCCAUCAGAAac **-3′** **3′-** GAGUUCUUACUUAUUAGGUAGUCUUUG **-5′**	NM_198933 NM_198934 NM_011137 NM_198932
Oct1.13.3	**5′-** GCAGUUUGCCAAGACUUUCAAACaa **-3′** **3′-** CUCGUCAAACGGUUCUGAAAGUUUGUU **-5′**	NM_198932 NM_011137 NM_198934 NM_198933

Upper case letters represent RNA bases, while lower case letters correspond to DNA bases.

Table I.4: crRNAs designed for CRISPR/Cas9 genome editing, their sequences, targets, and genomic location. The number, in brackets, of the CNS2 crRNAs corresponds to the drawing of their location described in figures I.8 B, I.9 B and SI.2 A.

crRNA name	Sequence 5′ → 3′	PAM	Target	Genomic location, GRCh38/hg38
CNS2_del_3′	TGTGTGCACCTAACATACCT	AGG	Outside CNS2, 3′ region	Chr12: 9,765,764 - 9,765,783
CNS2_del_5′ (1)	ACAAGTATTGATTAGTACCG	TGG	Outside CNS2, 5′ region	Chr12: 9,764,342 - 9,764,361
CNS2_I (2)	GGCTTAGAGAAATTAGTGCG	GGG	CNS2 region I	Chr12: 9,764,760 - 9,764,779
CNS2_N47 (3)	AGGAAAGCTGCATTGAGTTT	TGG	CNS2 region I, TSS -	Chr12: 9,764,814 - 9,764,833
CNS2_SRF (4)	TTCTACATATAAGGTCACAC	AGG	CNS2 region I, SRF BS	Chr12: 9,764,838 - 9,764,857
CNS2_N3 (5)	TCCAGGGTGAGACGTCAGAA	AGG	CNS2 region I	Chr12: 9,764,858 - 9,764,877
CNS2_Oct1_A (6)	TCAGAAAGGAAGTAAAGATG	AGG	CNS2 region II, Oct1 BS A	Chr12: 9,764,872 - 9,764,891
CNS2_Oct1_B (7)	CAAAGTGTGCAGAGAAGGTG	GGG	CNS2 region II, Oct1 BS B and TSS +	Chr12: 9,764,911 - 9,764,930
CNS2_N68 (8)	TATCAGACAGCTGCAGCAGC	AGG	CNS2 region II, GATA/MyoD BSs	Chr12: 9,764,945 - 9,764,964
CNS2_N27 (9)	CTGCAGCAGCAGGATGCTTT	ACC	CNS2 region II	Chr12: 9,764,955 - 9,764,974
CNS2_RUNX (10)	TGTCTTAGGTCGGAAGTCTG	TGG	CNS2 region III, RUNX BS	Chr12: 9,764,998 - 9,765,017
CNS2_III (11)	GGGTGAATTAGGTTTCTGAA	TGG	CNS2 region III	Chr12: 9,765,068 - 9,765,087

crRNA name	Sequence 5′ → 3′	PAM	Target	Genomic location, GRCh38/hg38
CNS2_IV_1 (12)	AATTTGCATTAATTTGAGAG	AGG	CNS2 region IV	Chr12: 9,765,140 - 9,765,159
CNS2_IV_2 (13)	AAAAACACAAAGGCACCTGC	AGG	CNS2 region IV	Chr12: 9,765,190 - 9,765,209
CNS2_N39 (14)	TTACCGTAAAGGGTCAATGA	TCC	CNS2 region IV-V	Chr12: 9,765,217 - 9,765,236
CNS2_N25 (15)	ACTTGAAGAGGACAGAAGGT	TGG	CNS2 region V, TSS +	Chr12: 9,765,284 - 9,765,303
CNS2_N29 (16)	TTTCAGATGTTCCATTGAAG	AGG	CNS2 region V	Chr12: 9,765,326 - 9,765,345

Table I.5: Neon Transfection System programming for cell electroporation.

Cell line	Pulse voltage (V)	Pulse width (ms)	Pulse number
Jurkat and Jurkat CRISPR-a/i	1325	10	3
C1R	1450	10	3
NK-92® MI	1200	10	4

Table I.6: crRNAs used for CRISPR/Cas9 deletion of CNS2 and its regulatory elements.

Region deleted	crRNAs used	Deletion length	Genomic location, GRCh38/hg38
CNS2	CNS2_del_3′ + CNS2_del_5′ (1)	1442 bp	Chr12: 9,764,342 - 9,765,783
Region I (ΔI)	CNS2_I (2) + CNS2_N3 (5)	98 bp	Chr12: 9,764,780 - 9,764,877
Region I + Oct1 A BS (ΔI+Oct1)	CNS2_I (2) + CNS2_Oct1_A (6)	112 bp	Chr12: 9,764,780 - 9,764,891
Oct1 A+B BSs (ΔCore E)	CNS2_N3 (5) + CNS2_Oct1_B (7)	53 bp	Chr12: 9,764,878 - 9,764,930
Oct1 A BS (ΔOct1 A)	CNS2_N3 (5) + CNS2_Oct1_A (6)	14 bp	Chr12: 9,764,878 - 9,764,892
Oct1 B BS (ΔOct1 B)	CNS2_Oct1_A (6) + CNS2_Oct1_B (7)	39 bp	Chr12: 9,764,892 - 9,764,930
Region 5′-II (Δ5′-II)	CNS2_Oct1_B (7) + CNS2_N27 (9)	44 bp	Chr12: 9,764,931 - 9,764,974
Region 5′-II + III (Δ5′II+III)	CNS2_Oct1_B (7) + CNS2_III (11)	157 bp	Chr12: 9,764,931 - 9,765,087
Region III (ΔIII)	CNS2_N27 (9) + CNS2_III (11)	113 bp	Chr12: 9,764,975 - 9,765,087

Table I.7: PCR primers used for CNS2 genotyping and the expected sizes of the amplicons. The number, in brackets, of the PCR primers corresponds to the drawing of their location described in **Figure SI.3 A**.

Primer	Sequence 5′ → 3′	Amplicon length		Genomic location, GRCh38/hg38
		Unedited	ΔCNS2	
CNS2_Del_Fw (1)	CTGTCCTAGCTGACACATGCT	1756 bp	314 bp	Chr12: 9,764,209 – 9,764,229
CNS2_Del_Rv (2)	GAAGAGACGTCTGTCTAGAATATG			Chr12: 9,765,941 – 9,765,964
I_II_III_srt_Fw (3)	AATGCCAGCTCACACTTTCC	806 bp	**ΔI:** 708 bp **ΔI+Oct1:** 694 bp **ΔOct1A:** 792 bp **ΔOct1B:** 767 bp **ΔCore E:** 753 bp **Δ5′-II:** 762 bp **Δ5′-II+III:** 649 bp **ΔIII:** 693 bp	Chr12: 9,765,362 – 9,765,381
I_II_III_Rv (4)	TGCAAGCAAACCCCTTTCAA			Chr12: 9,764,576 – 9,764,595

Table I.8: TaqMan qPCR probes targeting the whole CNS2 region.

Name	Sequence 5′ --> 3′	Amplicon length	Chromosome location, GRCh38/hg38
A_Fw	AAAGGAAAGTGTAGAGGGTCAAG	122 bp	Chr12: 9,764,408 - 9,764,529
A_Probe	ATGCAGTGTGGGAGAGATTTGGG		
A_Rv	CCATCAGGGCTAACCACAATAA		
B_Fw	CTTTCCTTCCTGTGTGACCTTAT	105 bp	Chr12: 9,764,827 - 9,764,931
B_Probe	TCCAGGGTGAGACGTCAGAAAGGA		
B_Rv	CCACCTTCTCTGCACACTTT		
C_Fw	CAGGATGCTTTTGGGTACATG	79 bp	Chr12: 9,764,964 - 9,765,042
C_Probe	GACCACAGACTTCCGACCTAAGACAATT		
C_Rv	AGATGAGCAGTTTGTCTCCG		
D_Fw	ACCGTAAAGGGTCAATGAAGG	77 bp	Chr12: 9,765,219 - 9,765,295
D_Probe	CCATGAGTAAACGGGGGATTTCCAGT		
D_Rv	AGGACAGAAGGTTGGAAAGTG		
E_Fw	TTTCCAACCTTCTGTCCTCTTC	134 bp	Chr12: 9,765,278 - 9,765,411
E_Probe	GCTTGGAAAGTGTGAGCTGGCATT		
E_Rv	ACCTTCACTGTCTCAAGTCAAA		
F_Fw	TCTGTGTGCACCTAACATACC	108 bp	Chr12: 9,765,762 - 9,765,869
F_Probe	AGGAAATGCCTCCAGGTGGATGAC		
F_Rv	GCCTACTTTCACCTCTCTCATAC		

Table I.9: TaqMan qPCR standard program.

Step	Temperature	Time (min:sec)	Cycles	Data collection
Polymerase activation	95°C	03:00	1	
Amplification:				
Denaturation	95°C	0:15	40	ON
Annealing/Extension	60°C	1:00		
Hold	4°C	∞	1	

Table I.10: SYBR green qPCR standard program.

Stage	Temperature	Time (min:sec)	Cycles	Data collection
Hold	95°C	02:00		
PCR	95°C	00:15	40	
	60°C	01:00		ON
Melting curve	95°C	00:15		
	60°C	01:00	Continuous	
	95°C	00:01		ON

Table I.11: SYBR green qPCR primers used for validation of DEGs

Target gene	Primer	Sequence 5′ → 3′
CCR1	**CCR1_Fw**	TCCTGCTGACGATTGACAGGTA
	CCR1_Rv	GTGCCCGCAAGGCAAAC
CCR2	**CCR2_Fw**	GCGTTTAATCACATTCGAGTGTTT
	CCR2_Rv	CCACTGGCAAATTAGGGAACAA
CCR5	**CCR5_Fw**	GCTGGTCATCCTCATCCTGATAA
	CCR5_Rv	ATGGCCAGGTTGAGCAGGTA
CCR8	**CCR8_Fw**	TGGCTGTTGTCCATGCCGTGTA
	CCR8_Rv	TGGGATGGTAGCCATAATGGCG
CXCR5	**CXCR5_Fw**	TGAAGTTCCGCAGTGACCTGTC
	CXCR5_Rv	GAGGTGGCATTCTCTGACTCAG
CCL3L3	**CCL3L3_Fw**	ACCGCCTGCTGCTTCAGCTAC
	CCL3L3_Rv	ATGACACTGGGCTTGGAGCACT
CCL4	**CCL4_Fw**	GCTTCCTCGCAACTTTGTGGTAG
	CCL4_Rv	GGTCATACACGTACTCCTGGAC
CCL4L1	**CCL4L1_Fw**	GCCTGCTGCTTTTCTTACACCG
	CCL4L1_Rv	GCAGACTTGCTTGCCTCTTTTGG

Results I

1. *In Silico* Characterization of Human and Mouse *CD69* Genomic *Locus* and its Regulatory Elements

A previous data-mining study of the CNSs regulatory features within *CD69 locus* proposed CNS2 as the putative main enhancer based on the enrichment of chromatin accessibility, modifications and bound TFs **[86]**. Given the availability of new ChIP-seq data, the TFBSs map was updated in the human *locus* using ChIP-seq data from human hematopoietic cell lines available from the ENCODE consortium and the ChIPBase v2.0 and ReMap databases (**Figure I.3 A**). We also mapped the bound TFs to the mouse *CD69 locus* using data from murine primary T and B cells available from the ChIPBase v2.0 and ChIP-Atlas databases (**Figure I.3 B**) to compare them. These maps showed that the human *CD69* 3' untranslated region (3'UTR), intron 1, promoter, CNS1, CNS2 and CNS4, and the murine 3'UTR, promoter, CNS1 and CNS2 have a higher density of TFBSs than the rest of noncoding sequences within the *CD69 locus*. Thus, the 3'UTR, promoter, CNS1 and CNS2 of both species contain numerous TFBSs, which suggest that they may play a particularly important role in the transcriptional regulation of *CD69* promoter. In agreement with previous observations, mouse CNS2 also shows the higher total number of TFBSs compared to the other regulatory regions within *CD69 locus*.

Previously, CNS2 in the mouse *CD69 locus* was subdivided into different subregions based on the presence of TFBSs conserved between six mammal species, and their contribution to the enhancer capacity of CNS2 was tested. However, that characterization left undefined regions where bound TFs have subsequently been identified and could play a relevant role. Moreover, considering that some regulatory mechanisms may have diverged during evolution and, that consequently, some important TFBSs might not be conserved, here CNS2 was subdivided into five regions based on the abovementioned updated ChIP-seq data. All bound TFs, the conserved ones as well as those bound to sites that are not conserved between six mammals species, were taken into account (**Figure I.4**)

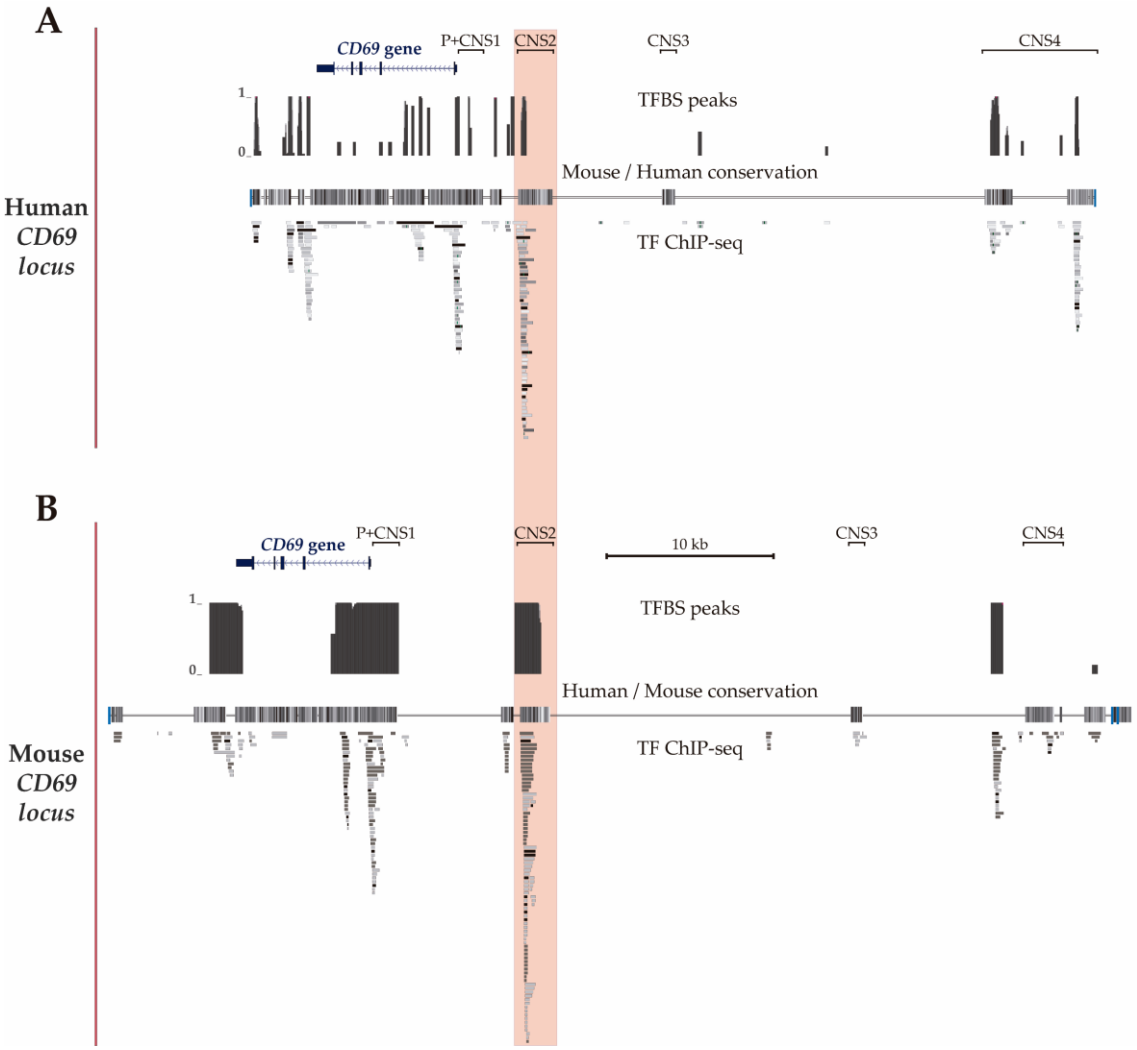


Figure I.3: *CD69* gene *locus*. UCSC genome browser showing the human **(A)** (GRCh37/hg19 assembly, chr12: 9,900,752-9,952,771) and mouse **(B)** (GRCm38/mm10 assembly, chr6: 129,259,289–129,321,308) CD69 *locus* and the four upstream conserved noncoding regions (CNS1-4). The alignment of the mouse and human genomes is displayed as a grayscale density plot that indicates the alignment quality, where darker values indicate higher levels of overall conservation. Transcription factor binding site peaks (TFBS peaks) show the proportion (between 0 and 1) of ChIP-seq data sets reporting TF binding at each location. The grey boxes represent different transcription factors found by ChIP-seq in human cell lines H1-hESC, A549, GM12878, HeLa-S3, IMR90, K562, HepG2, MCF-7, SK-N-SH, SK-N-SH_RA, HUVEC, AG04449, AG04450, AG09309, AG09319, AG10803, AoAF, BE2_C, BJ, Caco-2, Dnd41, ECC-1, Fibrobl, GM06990, GM08714, GM10847, GM12801, GM12864, GM12865, GM12872, M12873, GM12874, GM12875, GM12891, GM12892, GM15510, GM18505, GM18526, GM18951, GM19099, GM19193, GM19238, GM19239, GM19240, Gliobla, HA-sp, HAc, HBMEC, HCFaa, HCM, HCPEpiC, HCT-116, HEEpiC, HEK293, HEK293-T-REx, HFF, HFF-Myc, HL-60, HMEC, HMF, HPAF, HPF, HRE, HRPEpiC, HSMM, HSMMtube, HVMF, MCF10A-Er-Src, NB4, NH-A, NHDF-Ad, NHDF-neo, NHEK, NHLF, NT2-D1, Osteobl, PANC-1, PBDE, PBDEFetal, PFSK-1, ProgFib, RPTEC, Raji, SAEC, SH-SY5Y, SK-N-MC and T-47D (available at UCSC) and in mouse primary T CD4+, T CD8+ and B lymphocytes (TF-ChIP-seq, available at ChIP Atlas and ChIP-Base v2.0).

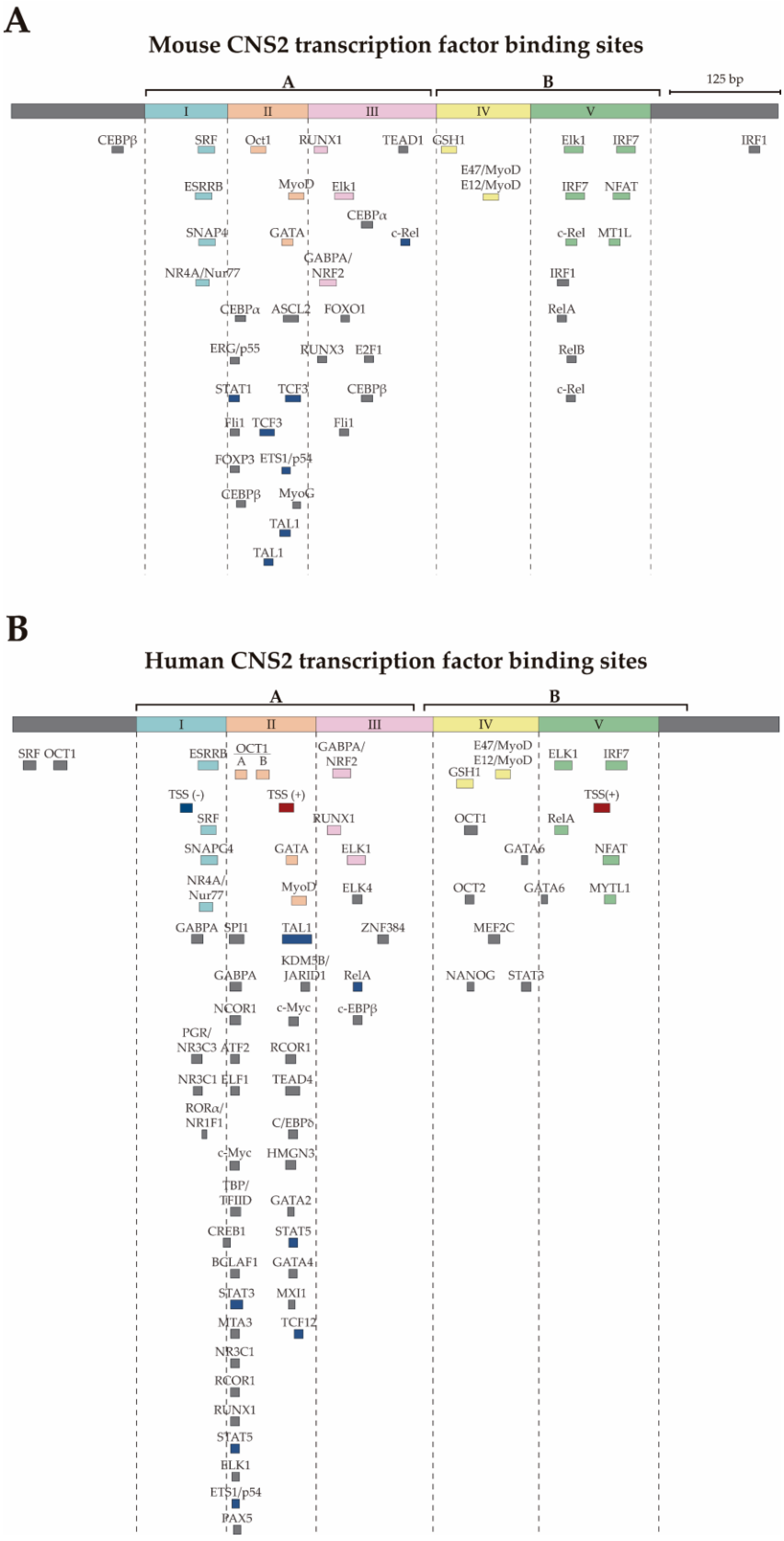


Figure I.4: ChIP-seq data shows an accumulation of TFBSs within mouse and human CNS2 regions I, II, and III. Map of the CNS2 regions and TFBS defined by ChIP-Seq available data. Color boxes indicate TFBSs conserved among six mammals species, as previously described [86]; grey boxes indicate non-conserved TFBSs, and dark blue boxes indicate non-conserved but shared TFBSs between mouse and human.

Transcription factors binding sites within the human CNS2 sequence were analyzed with the ChIP-seq data available at ChIPBase v2.0 and ReMap databases. As shown in **Figure I.4**, the conserved TFBSs that define the five regions of CNS2 are also depicted; moreover, comparison of bound TFs indicates an overall similarity in the TF distribution pattern between mouse and human CNS2 regions. The higher accumulation of TFBSs has been found in both mouse and human CNS2 regions I, II and III (A segment), whereas regions IV and V (B segment) exhibit a lower accumulation of described TFBSs. It is worth mentioning that detected TFBSs in human region II are more abundant than in mouse, even though in both species region II has most of the reported TFBSs because there are more experiments reported using human cells. The region II of both human and mouse share some activator TFBSs like the one for TAL1 **[104-106]**, ETS1/p54 **[107, 108]** or those for the STAT family of transcription factors **[109]**. Though interestingly, in human CNS2 region II have been detected binding sites for TFs described as transcriptional repressors such as MTA3 **[110, 111]**, NCOR1 **[112]** and RCOR1. Moreover, in both mouse and human region II it have been described TFBSs for the TCF repressor family of transcriptions factors; mouse TCF3 is thought to repress transcription in the absence of β-catenin **[113, 114]** and human TCF12 is thought to repress transcription through its interaction with Chd4 independently of the NuRD remodeling complex **[115, 116]**.

Several studies reported the posttranslational histone modifications in immune system cells. These modifications within the *CD69 locus* in naïve $CD4^+$ T cells are shown in **Figure I.5.** The *CD69* promoter is enriched in epigenetic marks of active regions, such as histone H3 lysine 4 monomethylation (H3K4me1), histone H3 lysine 4 trimethylation (H3K4me3) and histone H3 lysine 27 acetylation (H3K27ac), as well as in repressive epigenetic marks as histone H3 lysine 9 trimethylation (H3K9me3) and histone H3 lysine 27 trimethylation (H3K27me3). The CNS2 region contains repressive epigenetic marks such as H3K9me3 and H3K27me3, as well as epigenetic activation marks as H3K4me1, H3K4me3 and H3K27ac. Both the promoter and CNS2 shows an enrichment in RNA polymerase II and the histone H3 lysine 4 dimethylation (H3K4me2) epigenetic mark which, together with H3K27ac, define transcription factor binding regions (TFBRs) with high sensitivity.

Moreover, it has been described the enrichment of several histone acetyltransferases (HATs), such P300, MOF, PCAF and Tip60, as well as of several histone deacetylases (HDACs) such as HDAC1-6 at the CNS2 enhancer [117]. It is worth mentioning the special distribution of epigenetic marks in the region that we later characterized as the CNS2 Core E, where the repressive H4K27me3 is enriched while there is an absence of the H3K4me1, H3K4me3 and H3K27ac activation marks.

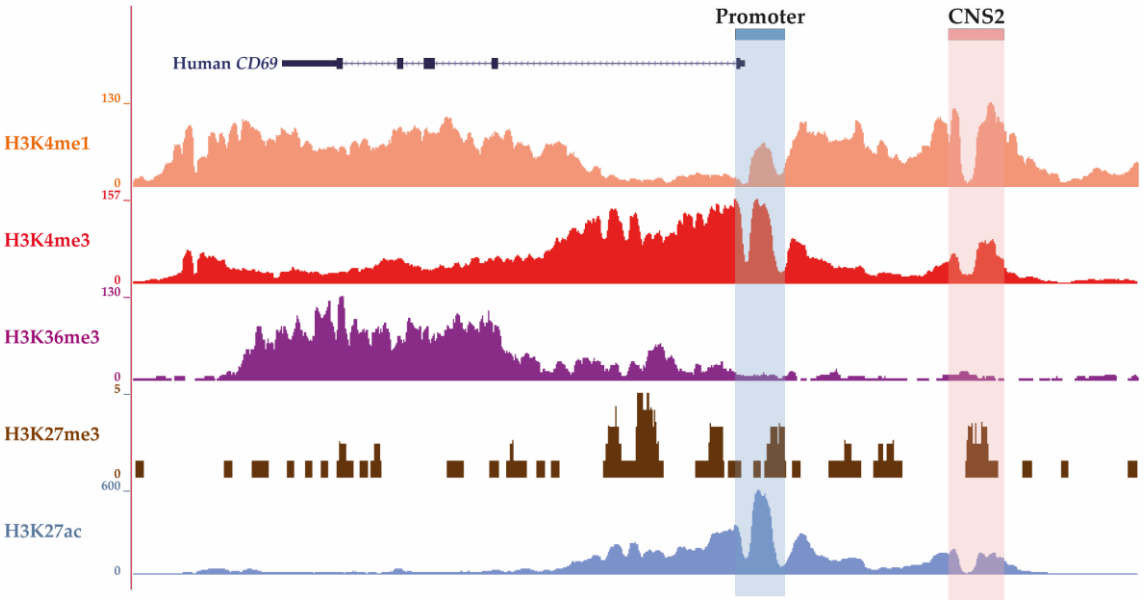


Figure I.5: Enrichment of posttranslational histone modifications within the human *CD69* locus. Histone posttranslational modifications in naïve CD4⁺ T cells were analyzed with the UCSC genome browser with the data publicly available at the Canadian Epigenetics, Environment and Health Consortium (CEEHRC) Network.

2. Assessment of the Transcriptional Enhancer Activity of CNS2 Regions and Its Enhancer Core by Luciferase Reporter Assays

First step was to assess the contribution of the different regions to mouse CNS2, whose high enhancer capacity was previously demonstrated by reporter assays. By transfecting reporter constructs in Jurkat T cells stimulated or not with PMA/Ionomycin (PMA/Io), it was compared the capacity of the complete mouse CNS2 sequence to enhance the mouse *CD69* promoter transcription with those of different CNS2 fragments: I+II+III, II+III and I+II (Figure I.6). Regions II+III have an inducible enhancer activity similar to that of regions I+II+III, while regions I+II have lower enhancer potential. Therefore, regions II+III seem to contain the most relevant enhancing features for the transcription of *CD69* promoter upon stimulation, while regions I, IV, and V have a more modest contribution.

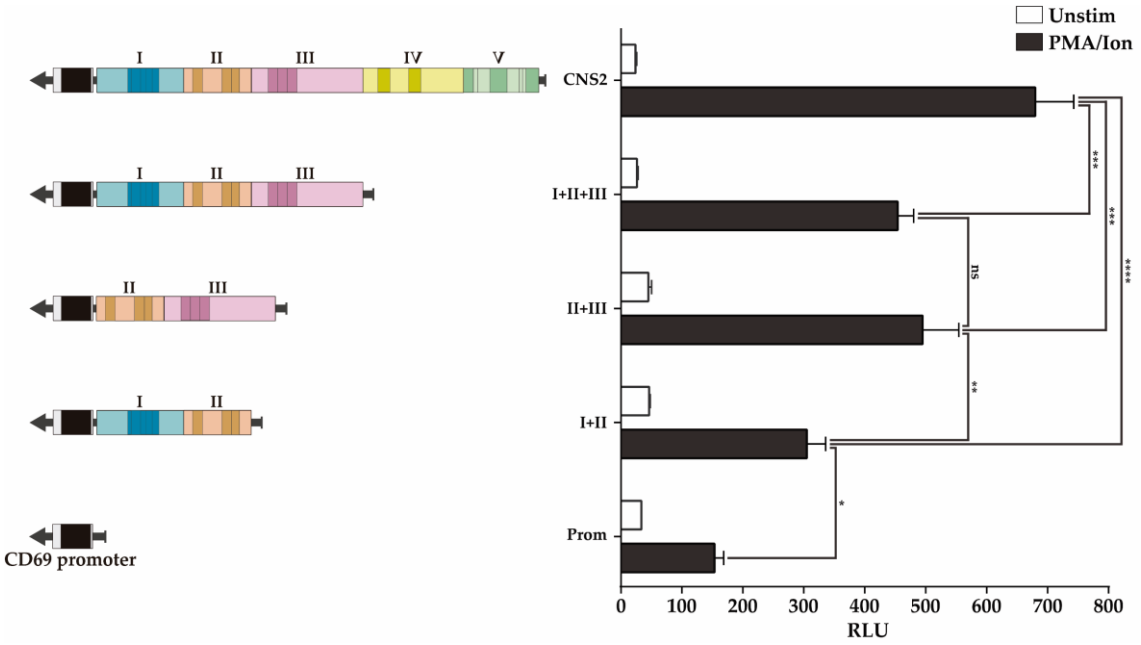


Figure I.6: Luciferase activity of CNS2 regions II+III is similar to that of the whole CNS2 sequence. (A) Map of the CNS2 regions and TFBSs defined by ChIP-Seq data. Color boxes indicate TFBSs conserved among six mammals species, as previously described [86]; grey boxes indicate non-conserved TFBSs identified by public ChIP-seq experiment databases. **(B)** Jurkat T cells were transfected with luciferase constructs carrying the mouse *CD69* promoter alone or together with the whole CNS2 or regions I+II, II+III, or I-III. The bars show the mean RLU of three independent experiments. The error bars show the standard deviation (SD). One-way ANOVA was applied; * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$). **RLU:** Relative luciferase units.

While region III contains three overlapping TFBSs, region II has two separate subregions: one with an Oct1 BS and another with GATA+MyoD TFBSs. The contribution of the individual regions I, II and III, as well as the subregions of region II separately or in combination with region I or III, was further assessed, but keeping the original sequence with the endogenous TFBSs distribution within the tested fragment.

Figure I.7 shows that region II significantly increases *CD69* promoter transcriptional activity upon stimulation, but neither region I nor region III increases *CD69* promoter transcriptional activity upon stimulation. The ~60 bp sequence of region II containing the Oct1 TFBS significantly increases the inducible enhancer activity in a manner similar to that in region II, while the ~65 bp sequence containing GATA+MyoD TFBSs did not significantly increases the inducible enhancer activity.

These results suggest that the subregion which contains the Oct1 TFBSs constitutes the most relevant regulatory feature of region II.

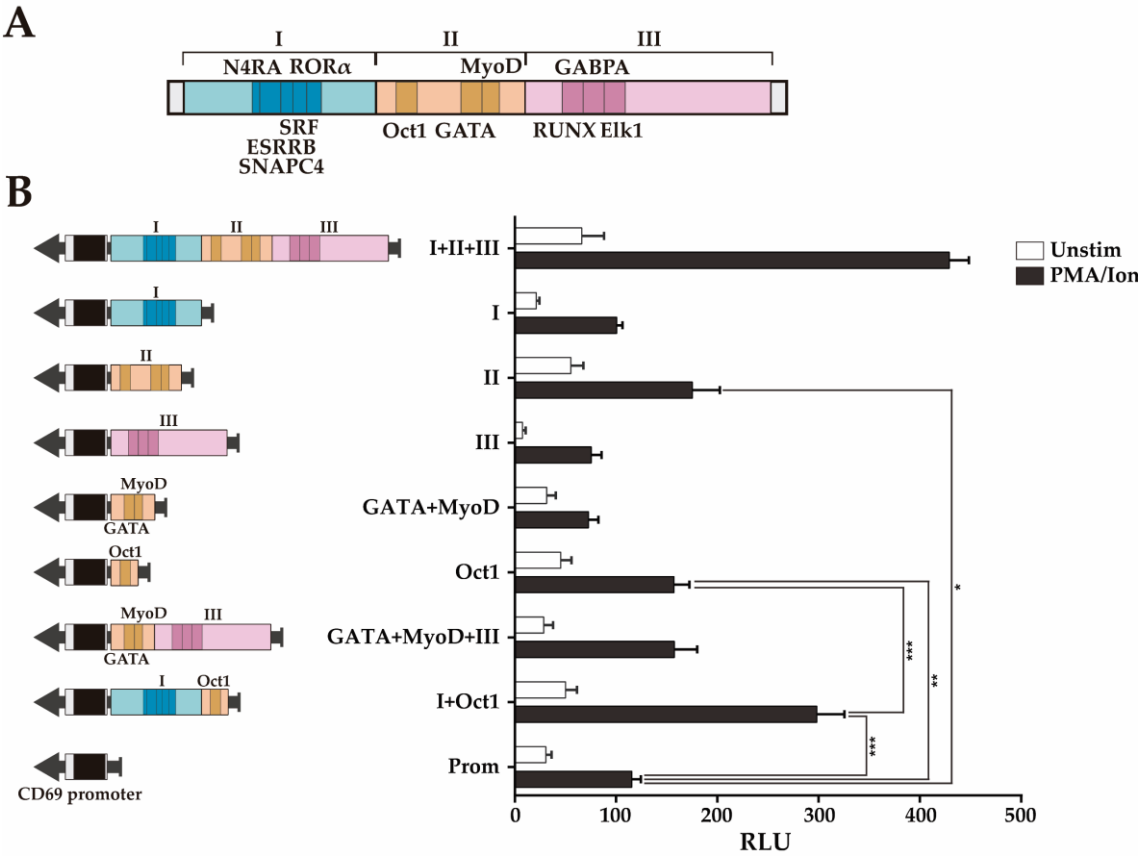


Figure I.7: Transcriptional enhancer activity of the CNS2 regions I+II+III and region II subregions. (A) Map of regions I-III and the conserved TFBSs within them. **(B)** Measurement of the luciferase activity after stimulation of Jurkat T cells transfected with constructs carrying the mouse *CD69* promoter alone or in combination with the different regions and subregions (I+II+III, I, II, III, GATA+MyoD, Oct1, GATA+MyoD+III, I+Oct1). The bars show the mean RLU. The error bars show the SD of three independent experiments with duplicate measurements. One factor ANOVA was applied to test for significant differences, with * ($p < 0.05$), ** ($p < 0.01$), and *** ($p < 0.001$). **RLU:** Relative luciferase units.

Moreover, considering that regions I and III alone do not show enhancing activity, these data indicate that the ~60 bp sequence containing the Oct1 TFBS is the shortest sequence identified within regions I+II+III that has transcriptional enhancer activity, likely acting as the enhancer Core of *CD69* transcription (Core E). Furthermore, the combination of this sequence with region I (~200 bp) had a transcriptional enhancing capacity as high as half of that of regions I-III, and similar to that regions I and II together (Figures I.6 and I.7). Although the combination of region III with the ~65 bp sequence GATA+MyoD (~190 bp) increased the transcriptional activity twofold compared to sequences containing the BS of each TF individually, it did not reach transcription levels significantly higher than those of the promoter alone.

Since GATA and MyoD BSs are between Oct1 and RUNX1, it is not possible to test Oct1+RUNX1 without altering the original TFBSs distribution. Thus, considering that regions II+III have almost as much enhancer activity as regions I-III (**Figure I.6**), it is suggested that the Oct1 TFBS seems to act as the enhancer Core of *CD69* transcription, and can synergize with either region I, region III or both to achieve a greater enhancement of transcription from *CD69* promoter. Altogether, these data suggest a cooperative modular model between the TFs in region II and those in regions I and III, being the 65 bp region II containing the Oct1 BS the Core enhancer of CNS2.

3. Recruitment of a Transcriptional Activator and Repressor to CNS2 TFBSs Suggests a Dual Regulatory Role for the CNS2 Core E

Targeted gene regulation has been proven to be a powerful strategy for interrogating, perturbing and engineering cellular systems **[23, 118-123]**. So next, it was assessed how the recruitment of a transcriptional activator or repressor to different regions of CNS2 enhancer might drive *CD69* gene regulation, measured as a change in cell surface CD69 protein expression. First, the effect of a potent transcriptional activator recruitment to different CNS2 regions was analyzed in steady state T cells. Several gRNAs were introduced in a Jurkat T cell line that constitutively expresses a death Cas9 fused to the VP64 transcriptional activator (dCas9-VP64 Jurkat CRISPR-a) **[18]**.

The measurement of cell surface CD69 protein expression 24 hours after single or combined gRNA transfection shown a high increase in several regions of CNS2 (**Figure I.8**). The highest activation with a single gRNA was observed when the VP-64 transcriptional activator is recruited to the CNS2 Core E, thus pointing to this sequence as the major area containing TFs that acts as repressors for CD69 transcription. Recruitment of the VP-64 activator to the Core E increases *CD69* expression by fivefold with gRNA 6, which targeted the Oct1 conserved TFBS, and threefold when the activator is recruited to the 3' or 5' adjacent sequences with the gRNAs 5 and 7, respectively. On the contrary, recruitment of the VP64 activator to the regions III and IV with the gRNAs 11, 12, 13 and 14 defined an area with little effect, and its recruitment to the region III with the gRNA 10 has no effect on CD69 protein expression.

Nevertheless, the activator increased *CD69* expression approximately twofold when was recruited to region I (gRNAs 2, 3 and 4), and region V (gRNAs 15 and 16).

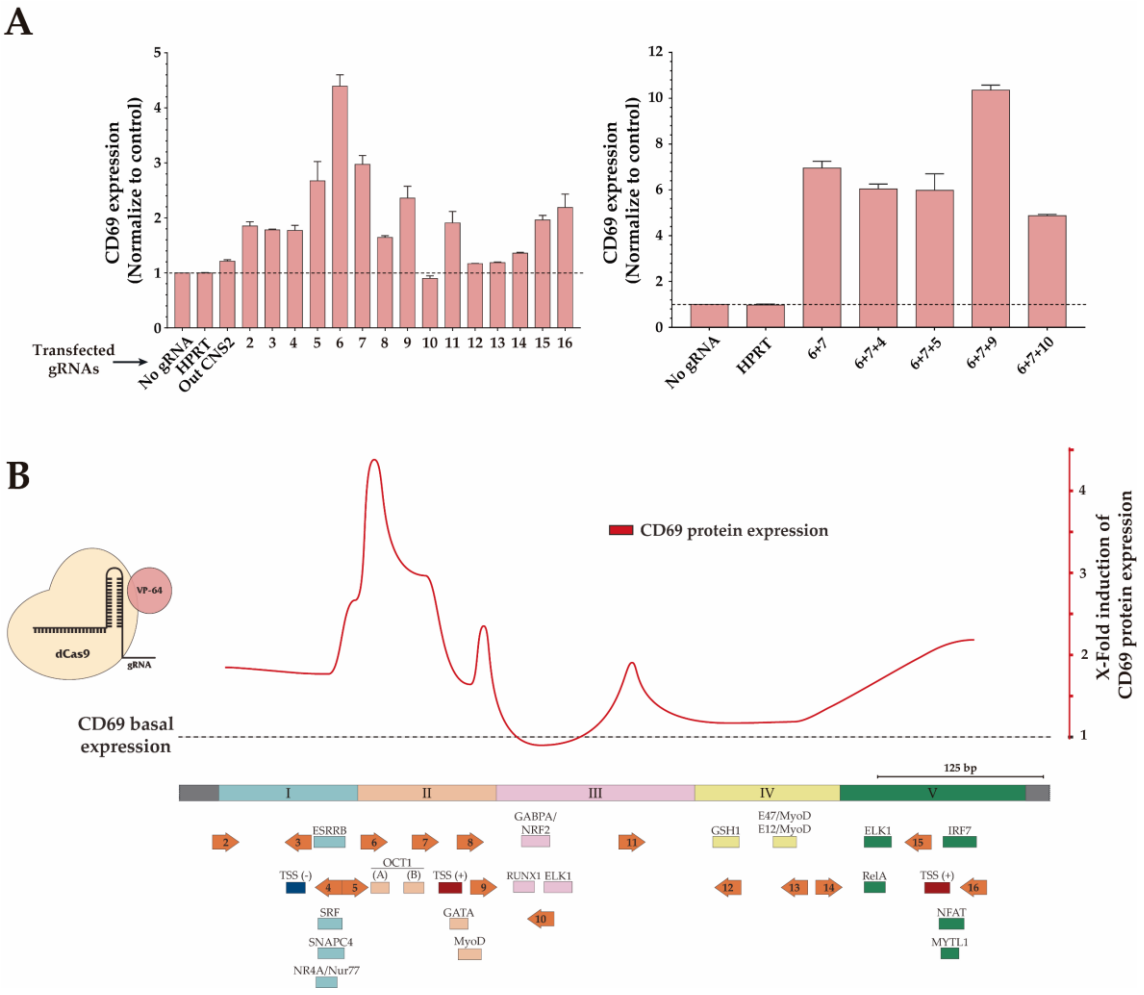


Figure I.8: Recruitment of the dCas9-VP64 fusion protein with gRNAs targeting CNS2 sequences supports the role of the Core E as a potent regulator of *CD69* expression. A) CD69 expression 24 hours after single (left panel) or combined (right panel) gRNA transfection, presented as a mean ± SD of at least three independent experiments**. B)** Schematic representation (red line) of *CD69* protein expression pattern above each tested gRNA (orange arrows).

Combination of the two gRNAs targeting the Core E (gRNAs 6+7) increases CD69 protein expression sevenfold. Combination of gRNAs 6 and 7 together with gRNA 4, 5 (targeting the region I) or with gRNA 10 (targeting the region III) did not increase *CD69* expression to a greater extent than combined gRNAs 6 and 7. Combination of three gRNAs of the four that displayed the highest transcriptional activation (gRNAs 6+7+9) increases CD69 protein expression tenfold, suggesting a synergic activation effect.

A similar expression pattern was observed at 6 and 3 hours upon transfection of single or combined gRNAs (**Figure SI.2**). These results suggest the possibility that a functional repressor module would be placed at the Core E in steady state T cells as part of a preactivated state of *CD69* transcription.

Next, it was analyzed the effect of the recruitment of a transcriptional repressor to the CNS2 regions by gRNAs in T cells that constitutively expresses a death Cas9 fused to the KRAB transcriptional repressor (dCas9-KRAB Jurkat CRISPR-i). CD69 protein expression was measured 18 hours after gRNAs transfection plus 6 additional hours of PMA/Io stimulation. By these settings, the decrease in CD69 expression will allow us to define the areas within CNS2 where the KRAB transcriptional repressor displaces the activation transcription factors (**Figure I.9**). The recruitment of the KRAB repressor to the CNS2 Core E region with gRNAs 5, 6 and 7 decreases ~50% CD69 protein expression, indicating that the Core E binds activator TFs upon *CD69* induction in T cells. However, the highest repression of *CD69* expression, reaching more than 80%, was achieved by recruitment of the repressor to the 5' region adjacent to the Core E with the gRNAs 9 and 10. The transcriptional repressor slightly decreases CD69 expression when was recruited to region I (gRNAs 2, 3 and 4) and region IV (gRNAs 12 and 13). At steady state, *CD69* expression changes were not observed (data not shown).

Together, these data suggest that, at the steady state, the CNS2 Core E may be occupied with a set of repressive transcription factors or cofactors which, upon stimulation, may be changed by the recruitment of the transcriptional activation machinery both to the Core E and its 5' adjacent region (5'C region), being the 5'C region the major platform for the enhancer activity of *CD69* gene transcription. The Core E itself behaves as a dual regulator of *CD69* expression depending on the cell activation context.

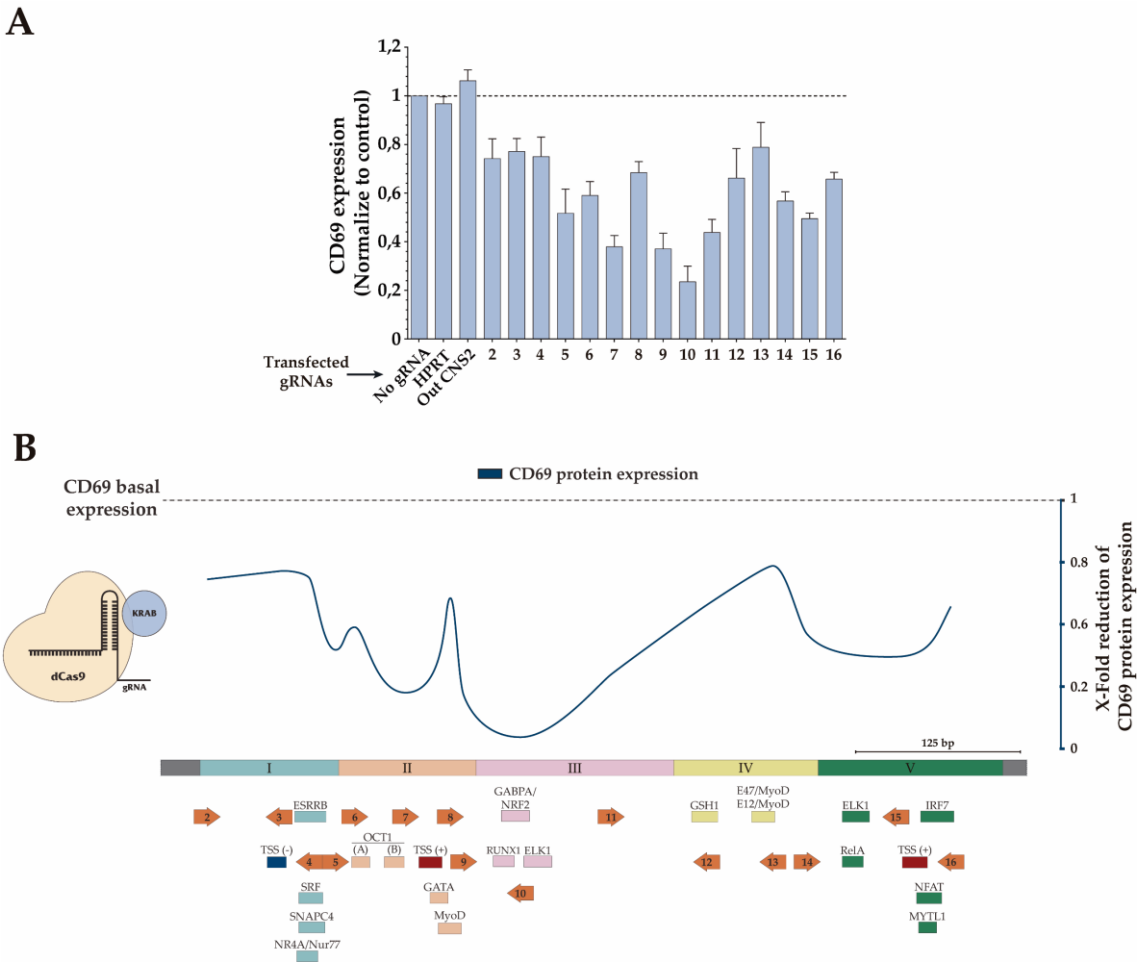


Figure I.9: Recruitment of dCas9-KRAB transcriptional repressor with gRNAs targeting CNS2 sequences indicates that region III may behave as a potent activator upon stimulation. A) Human T cells that constitutively express the dCas9-KRAB fusion protein were transfected with gRNAs (orange arrows) targeting the main TFBSs within CNS2. Eighteen hours later cells were stimulated with PMA/Io for 6 hours and CD69 protein expression was measured. Data are presented as a mean ± SD of at least three independent experiments. **B)** Schematic representation (blue line) of CD69 protein expression pattern above each tested gRNA (orange arrows).

4. CD69 Expression is Differentially Affected by the Deletion of the CNS2 Core E and Its Adjacent Sequences

Next, we aimed to study the contribution of each CNS2 region in its genomic context, deleting them by the CRISPR/Cas9 genome editing system. CD69 protein expression of unedited Jurkat T cells and the selected clones for each deleted region was analyzed 24 hours after PMA/Io stimulation. It was found that upon stimulation, the deletion of the entire CNS2 Core E (ΔCore E) or its parts (ΔOct1 A and ΔOct1 B) led to an important 230%, 160% and 200%, respectively, increase of CD69 expression, thus revealing its repressor function during activation.

Deletion of the CNS2 region I (ΔI) led to a 60% decrease of CD69 expression after stimulation, reaching the 70% decrease when the adjacent TFBSs cluster of region II was also deleted (ΔI+Oct1 A). Deletion of ~65 bp of the 5′ sequence of region II (Δ5′-II), adjacent to the Core E, led to an impairment on CD69 expression of ~80% after stimulation; while deletion of the region III (ΔIII) led to an impairment of ~85% after stimulation (**Figure I.10 A**). Accordingly, deletion of the 5′ sequence of region II together with region III (Δ5′C, ~160 bp) led to an impairment of CD69 expression of ~95% after stimulation. These results suggest that the ~160 bp 5′C region exerts an important role in *CD69* transcriptional activation regulation.

The dual regulatory role of the Core E was confirmed by its deletion in the Jurkat dCas9-VP64 (Jurkat CRISPR-a) cell line, which led to a ~fourfold increment in CD69 protein expression both at rest (data not shown) and 24 hours after PMA/Io stimulation (**Figure I.10 B**). Furthermore, deletion of the whole CNS2 sequence led to a failure in CD69 expression after stimulation, suggesting that CNS2 would be a necessary element for CD69 expression, which is in agreement with a similar observation by Redondo et al. in unmodified Jurkat T cells (manuscript in preparation).

Overall, these results highlight a global picture of CNS2 in which the 3′ and 5′ extended Cores are essential to enhance *CD69* transcription, acting together with adjacent TFBSs to achieve its maximum enhancing properties. In this context, the Core E comprising the Oct1 BSs would act as a repressor, as its deletion leads to a CD69 overexpression.

5. Characterization by Chromatin Immunoprecipitation of Major CNS2 Transcription Factors

Among the TFs described bound to CNS2 that are conserved between species, Chd4, Oct1 and RUNX1 have been described as activators or repressor depending on the genomic context and the cofactors that binds to them **[124-131]**. The presence in CNS2 of these TFs and SPI1 as a possible coactivator of RUNX1 was assessed.

Figure I.11 depicts the quantification of the bound TFs assessed in unstimulated T cells and 24 hours after PMA/Io stimulation. At steady state, RUNX1 and Oct1 showed a similar pattern of binding with a higher enrichment in region III and lower in region V, whereas Chd4 slightly bind to the same regions (**Figure I.11**, dashed line).

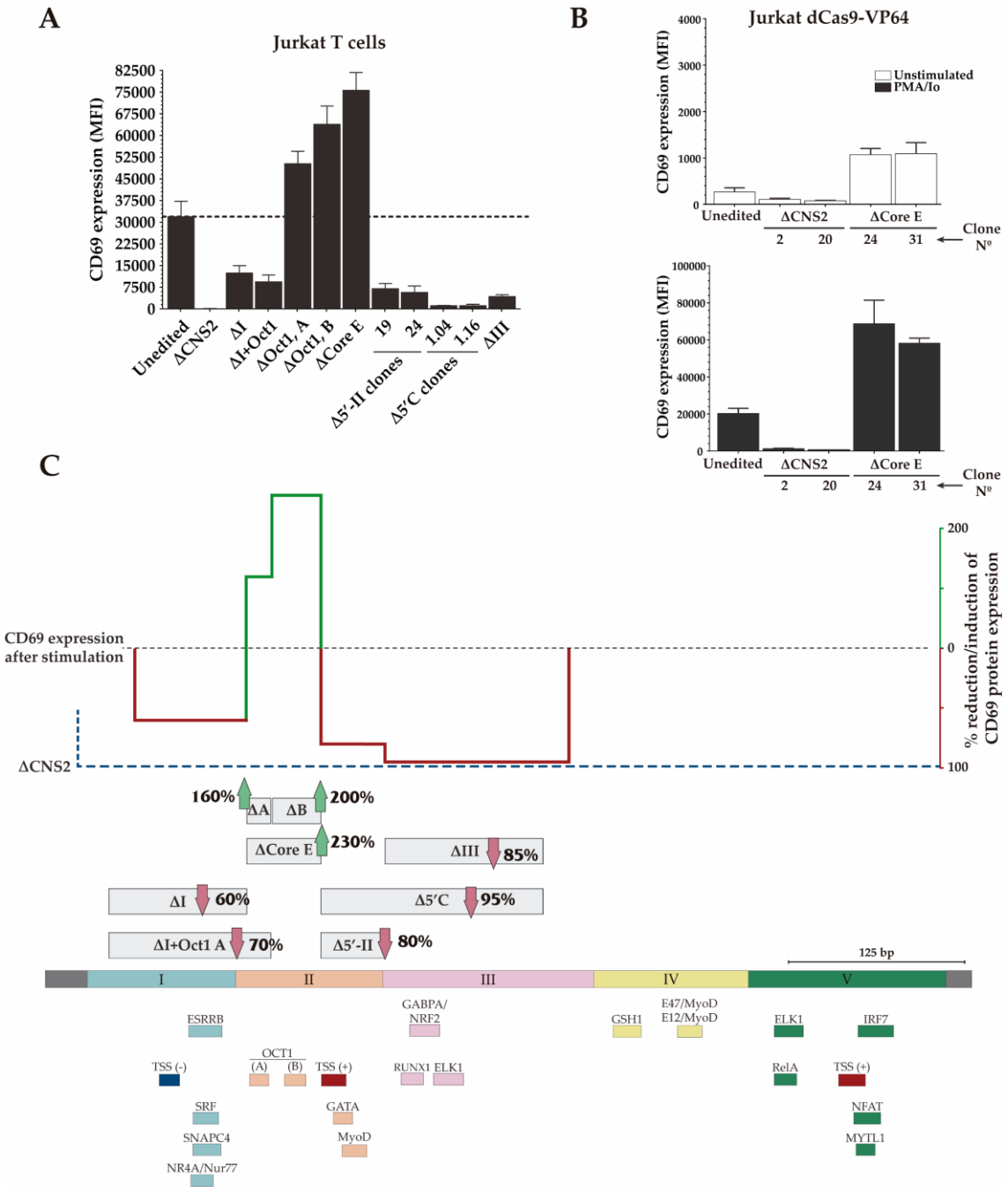


Figure I.10: CRISPR/Cas9 deletion of the human TFBSs contained in regions I-III of CNS2 reveals a suppressive function of the region II containing the Core E. A) Mean fluorescence intensity (MFI) of CD69 protein expression on unedited Jurkat T cells and homozygous clones for the deletion of the whole CNS2 fragment or different sequences containing TFBSs of CNS2. **B)** CD69 expression of Jurkat dCas9-VP64 lacking the whole CNS2 sequence or the Core E region was measured in unstimulated conditions (upper panel) or 24 hours after PMA/Io stimulation (bottom panel) **C)** Schematic representation of human CNS2 regions I, II and III, the conserved TFBSs (color boxes) and the length of every CRISPR/Cas9 deleted sequence (grey boxes). Red arrows indicate the % reduction of CD69 expression after stimulation, whereas green arrows indicate the % increase of CD69 expression after stimulation, compared to the unedited Jurkat T cells. Data are presented as a mean ± SD of at least three independent experiments for every deleted region.

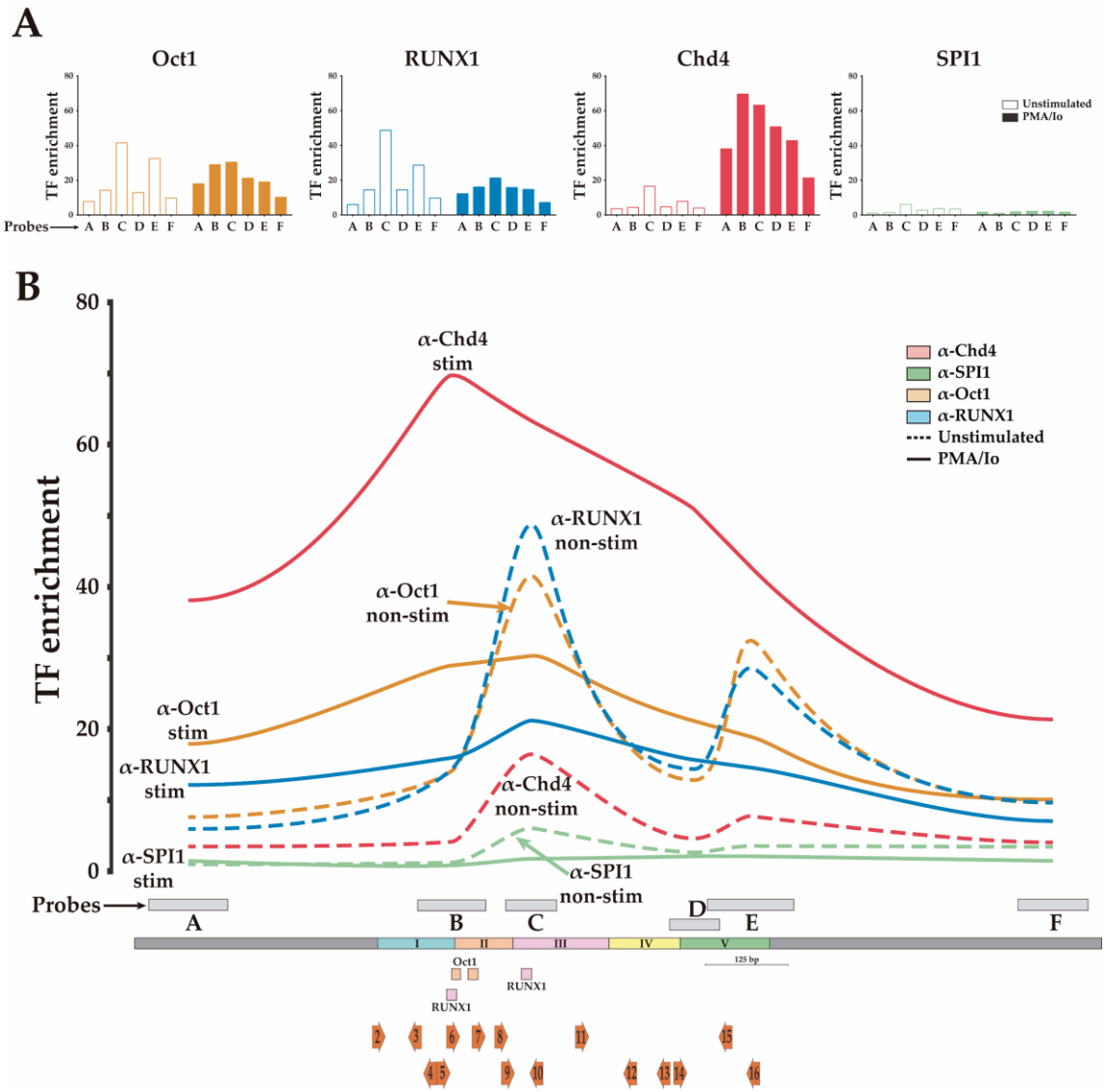


Figure I.11: Chromatin immunoprecipitation of CNS2 shows an enrichment of RUNX1 and Oct1 in unstimulated conditions, whereas after stimulation Chd4 increases its binding to CNS2. A) Oct1, RUNX1, SPI1 and Chd4 enrichment in unstimulated conditions (unfilled bars) or 24 hours after PMA/Io stimulation (filled bars), within the CNS2 regions analyzed (probes 1-6). Data for Oct1, RUNX1, and Chd4 are presented as a representative experiment of a duplicate, whereas data for SPI1 represent one experiment**. B)** Schematic representation of Oct1 and RUNX1 (left panel), and SP1 and Chd4 (right panel) enrichment within the CNS2 sequence. qPCR probes (A-F) are depicted as grey squares, and orange arrows represent the gRNAs used in previous experiments. Dashed lines represent the TF enrichment in unstimulated conditions, whereas continuous lines represent the TF enrichment 24 hours after PMA/Io stimulation.

Upon stimulation, Oct1 shift its enrichment to the Core E region and Chd4 dramatically increases its enrichment in the Core E region, whereas RUNX1 enrichment decreases along the whole CNS2 sequence. SPI1 did not bind to CNS2 either at resting conditions or after stimulation, therefore it does not have a relevant role in the transcriptional regulation of *CD69* in T cells.

These data may indicate that Oct1 and Chd4 form part of the same complex in the Core E, formed 24 hours after stimulation, whereas Oct1 and RUNX1 seem to be enriched coordinately outside of the Core E in the CNS2 putative poised steady state. Experiments deeper addressing the binding dynamics of these TFs are needed for further decipher their functional interaction in T cells.

6. Deletion of the CNS2 Core E Leads to a Dysregulation of the Cytokine-Cytokine Receptor Gene Interaction Network in T Cells

Due to the importance of the Core E, as demonstrated by its transcriptional activation, repression and deletion, we decided to perform an RNA-seq analysis of ΔCore E Jurkat T cells at resting and 3 hours after stimulation to assess whether the Core E affected the expressed gene pool.

In unstimulated conditions, of a total of 15,248 genes, ΔCore E T cells displayed 696 differentially expressed genes (DEGs), of which 379 were upregulated and 317 downregulated. After 3 hours of PMA/Io stimulation, of a total of 15,081 genes, only 165 were DEGs, of which 64 were upregulated, and 101 were downregulated (**Figure I.12 A** and **B**). Dysregulated genes (Fold Change ≥ 2, P-value < 0.05, False Discovery Rate < 0.01) are shown in the volcano plot (**Figure I.12 C**).

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway classification analysis of DEGs revealed that both at resting and after stimulation, ΔCore E dysregulated genes are involved mainly in signal transduction and are immune system related genes. KEGG pathway enrichment showed that in unstimulated conditions DEGs are involved in cytokine-cytokine receptor interaction and are related to the ABC family of transporters, whereas after stimulation dysregulated genes are involved mainly in cytokine-cytokine receptor interaction and Th1/Th2 cell differentiation (**Figure I.13**).

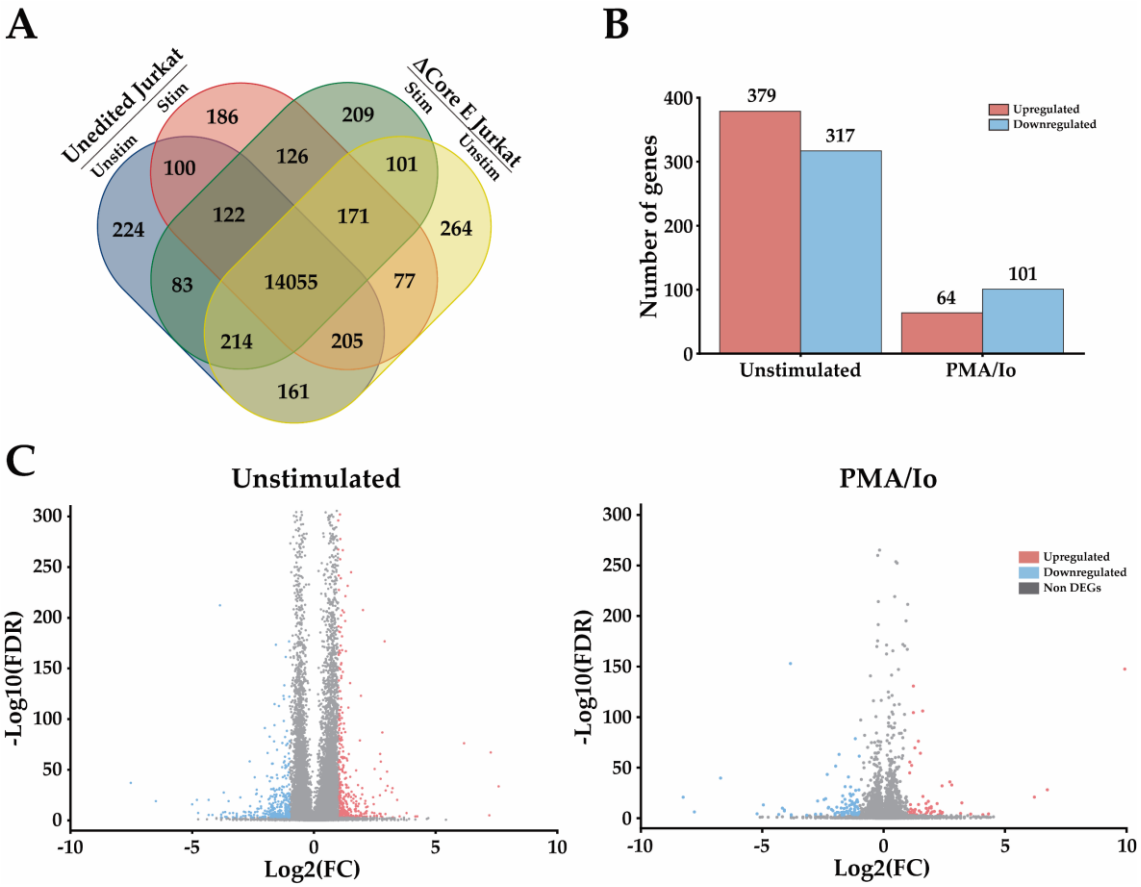


Figure I.12: RNA-seq of ΔCore E Jurkat T cells, which overexpress CD69. A) Venn diagram of differentially expressed genes (DEGs) between unstimulated or PMA/Io stimulated ΔCore E and unedited Jurkat T cells. **B)** Number of DEGs between ΔCore E Jurkat T cells unstimulated or 3 hours after PMA/Io stimulation, compared to the unedited Jurkat T cells**. C)** Volcano plot of unstimulated (left panel) and PMA/Io stimulated (right panel) cells, filtered out the dysregulated genes (FC≥2, P<0.05, FDR<0.01). Red and blue dots represent upregulated and downregulated genes, respectively.

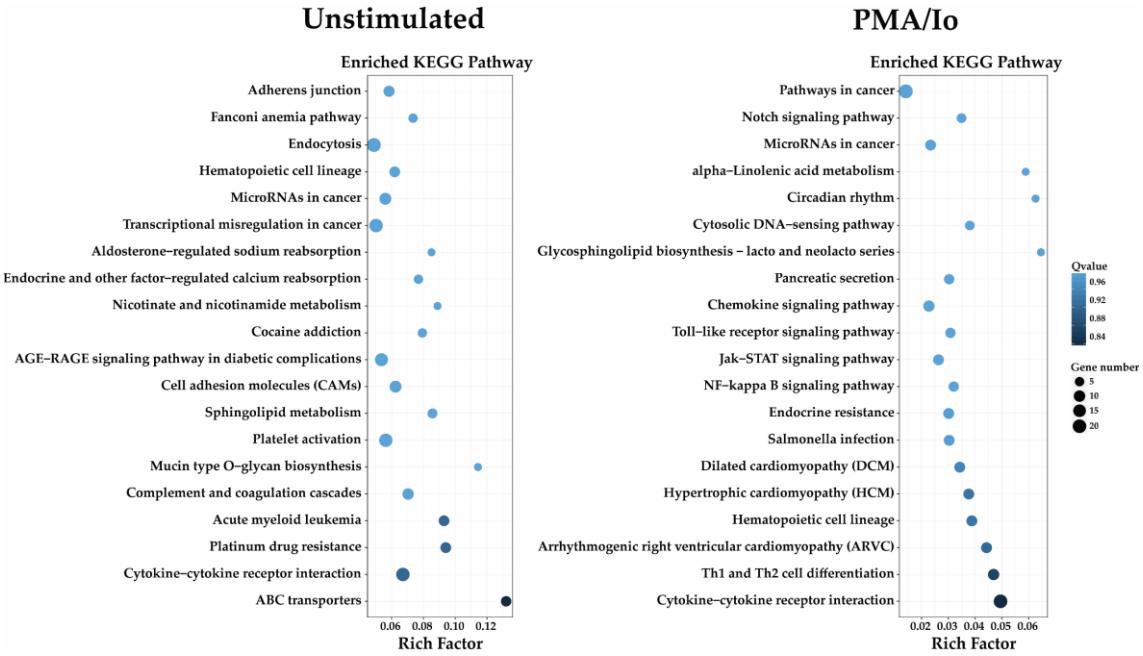


Figure I.13: KEGG functional pathway enrichment analysis of DEGs. Kyoto Encyclopedia of Genes and Genomes (KEGG) functional pathway analysis of DEGs between unedited or ΔCore E Jurkat T cells in unstimulated conditions (left panel) or after PMA/Io stimulation (right panel).

STRING analysis of upregulated genes shows an enrichment of the cytokine-cytokine receptor network in the ΔCore E Jurkat T cells. qPCR analysis of chemokine receptors confirms the upregulation of *CCR1*, *CCR2*, and *CCR5* in unstimulated conditions, whereas their ligands *CCL3L3*, *CCL4* and *CCL4L1* were upregulated upon stimulation in ΔCore E Jurkat T cells (**Figure I.14**). These observations open up the possibility that CD69 overexpression observed in cells lacking the Core E has a direct impact on cytokines-cytokine receptors regulation. However, as human *CCR1*, *CCR2* and *CCR5* are located in the same gene cluster at chromosome 3, and *CCL3L3*, *CCL4* and *CCL4L1* are located in the same gene cluster at chromosome 17, we cannot rule out an inter-chromosomal interaction between the Core E and the abovementioned gene *loci*.

Overall, these results pave the way for the understanding of the fine-tuned *CD69* expression regulation by the CNS2 regulatory Core E in T cells, highlighting the importance of *CD69* in T cell physiology.

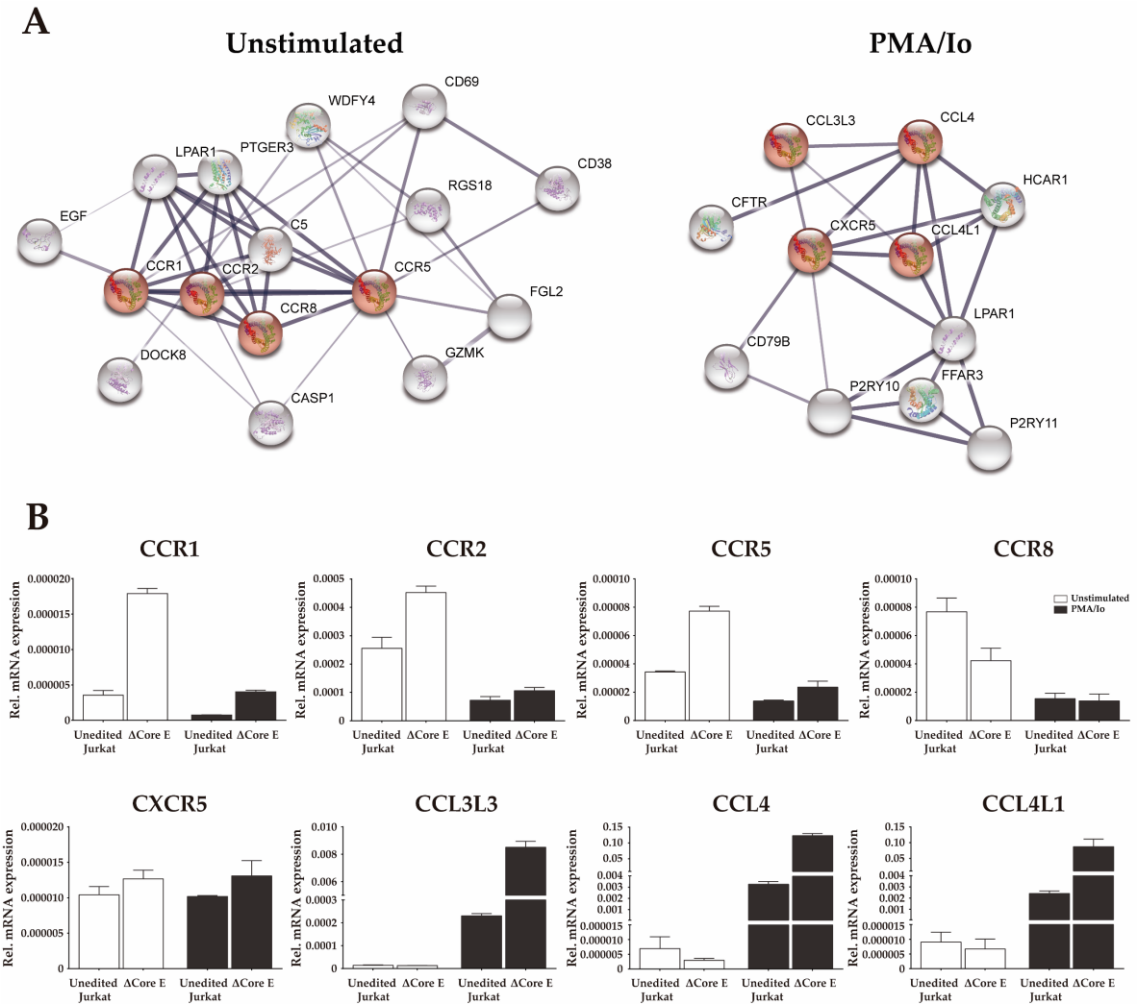


Figure I.14: STRING analysis of upregulated genes shows an enrichment of cytokine-cytokine receptor interaction related genes in ΔCore E Jurkat T cells. A) Search tool for recurring instances of neighboring genes (STRING, available at string-db.com **[101]**) analysis defined an association network of cytokine-cytokine receptor signal pathway (FDR 3.7×10^{-6}), represented as red circle genes**. B)** qPCR analysis of DEGs related to cytokine-cytokine receptor signaling confirmed the upregulation of *CCR1*, *CCR2*, and *CCR5* in unstimulated ΔCore E Jurkat T cells, whereas after stimulation their ligands *CCL3L3*, *CCL4* and *CCL4L1* were upregulated. Data are presented as a mean ± SD of two independent experiments.

Discussion I

CD69 has a basal expression at steady state, being quickly induced upon stimulation, and has been described as a poised gene **[117].** In this work, *CD69* enhancer transcriptional regulatory landscape has been characterized using the CRISPR/Cas9 toolbox, which has been proven to be an efficient approach in combination with classical methodologies. We found that *CD69* expression is controlled by CNS2 through a Core region (Core E) with a dual role as a transcriptional repressor and activator, and a 5' region adjacent to the Core E (5'C region) where the major transcriptional machinery is associated, with an activation function.

The recruitment of the dCas9/VP-64 transcriptional activator **[17, 18, 20, 132, 133]** to non-coding regulatory sequences within a gene *locus* may lead to an increase in its transcriptional induction due to the VP-64 recruitment of basal transcription machinery that displaced the already bound transcription factors, thus identifying stimulation-responsive elements (CaREs) in unstimulated cells **[18]**. Complementarily, the similar recruitment of the dCas9-KRAB transcriptional repressor **[20, 122, 132-134]** in stimulated cells leads to an impairment in the gene transcriptional induction when KRAB binding displaces the activating TF machinery.

Since the recruitment of the VP-64 to the CNS2 Core E in steady state led to a high induction of *CD69* protein expression, whereas the recruitment of the KRAB to the Core E upon stimulation led to ~50% loss of *CD69* expression, the Core E may act as a switch of transcription repressing the *CD69* gene *locus* at basal state but helping to the transcriptional activation machinery during *CD69* gene induction. In addition, CRISPR/Cas9 deletion of the Core E led to an augmented CD69 expression both at steady state and upon stimulation, suggesting that at both stages the Core E is a module with an equilibrium of activator and repressive TFs, probably helping to hold the transcriptional activity. Moreover, the ENCODE data of naïve $CD4^+$ T cells shows that the repressive epigenetic mark H4K27me3 is enriched in the Core E, while there is an absence of the H3K4me1, H3K4me3 and H3K27ac activation marks, though the surrounding CNS2 sequences are enriched in the H3K4me1, H3K4me3 and H3K27ac activation marks.

It is tempting to speculate that, at basal state, the Core E repressive function contributes to the previously defined poised state of the *CD69* gene and that, upon stimulation, the Core E may behave controlling CD69 overexpression (**Figure I.15**).

On the other hand, the recruitment of the KRAB repressor to the ~160bp 5' region adjacent to the Core E (5'C region) and its deletion almost completely impaired the *CD69* expression induced by activation (~90% of decrease), therefore delineating this region as a major platform where activator and co-activator TFs are recruited upon stimulation. These results define the regulatory dynamics of CNS2 as exposed in **Figure I.15**. At steady state, a set of repressive TFs may be bound to the CNS2 Core E, acting as a brake for *CD69* transcription, whereas upon stimulation there may be a recruitment of the TF machinery at both the Core E and the 5'C region that act as a potent enhancer of *CD69* transcription. We propose that activation TFs that bind to the CNS2 Core E after a long period of stimulation may act modulating the enhancer function of those bound to the 5'C region, controlling *CD69* transcriptional induction.

Occupancy by specific TFs has been described to predict enhancer activity of a given region **[106]**. *In silico* analysis of CNS2-bound TFs in hematopoietic cells shows an accumulation of TFBSs within regions I, II, and III, as well as a great enrichment in H3K4me2 epigenetic mark which, together with H3K27ac, define TFBRs with high sensitivity **[135]**. In this work, our ChIP assays performed in human T cells showed an enrichment of Oct1 and RUNX1 at steady state in the CNS2 5'C region, whereas upon stimulation RUNX1 enrichment highly decreased and Oct1 remained bound, but shifting to the Core E region. Therefore, it may be consider the possibility that Oct1 and RUNX1 may play an important function in the Core E regulation since these factors are known to have both activating **[136-138]** and repressive activity **[139-142]**. Moreover, these TFs have been described to play a role in $CD4^+$ T cell differentiation **[127]** and memory formation **[128]**, orchestrating the fate and function of $CD4^+$ effector T cells as a switchable stabilizer of the repressed an inducible states **[129]**, acting at hypersensitivity sites that are distant from promoters of target genes in several cases.

Since it has been reported the formation of a complex between RUNX and Oct1 that regulates the expression of several genes [143], it may be considered the possibility that the interaction of these TFs may regulate the enhancer function of CNS2. This assumption is supported by the presented results in which inhibition of RUNX1 and Oct1 TFs with interfering RNAs synergistically led to a reduction of CD69 expression (Figure SI.5).

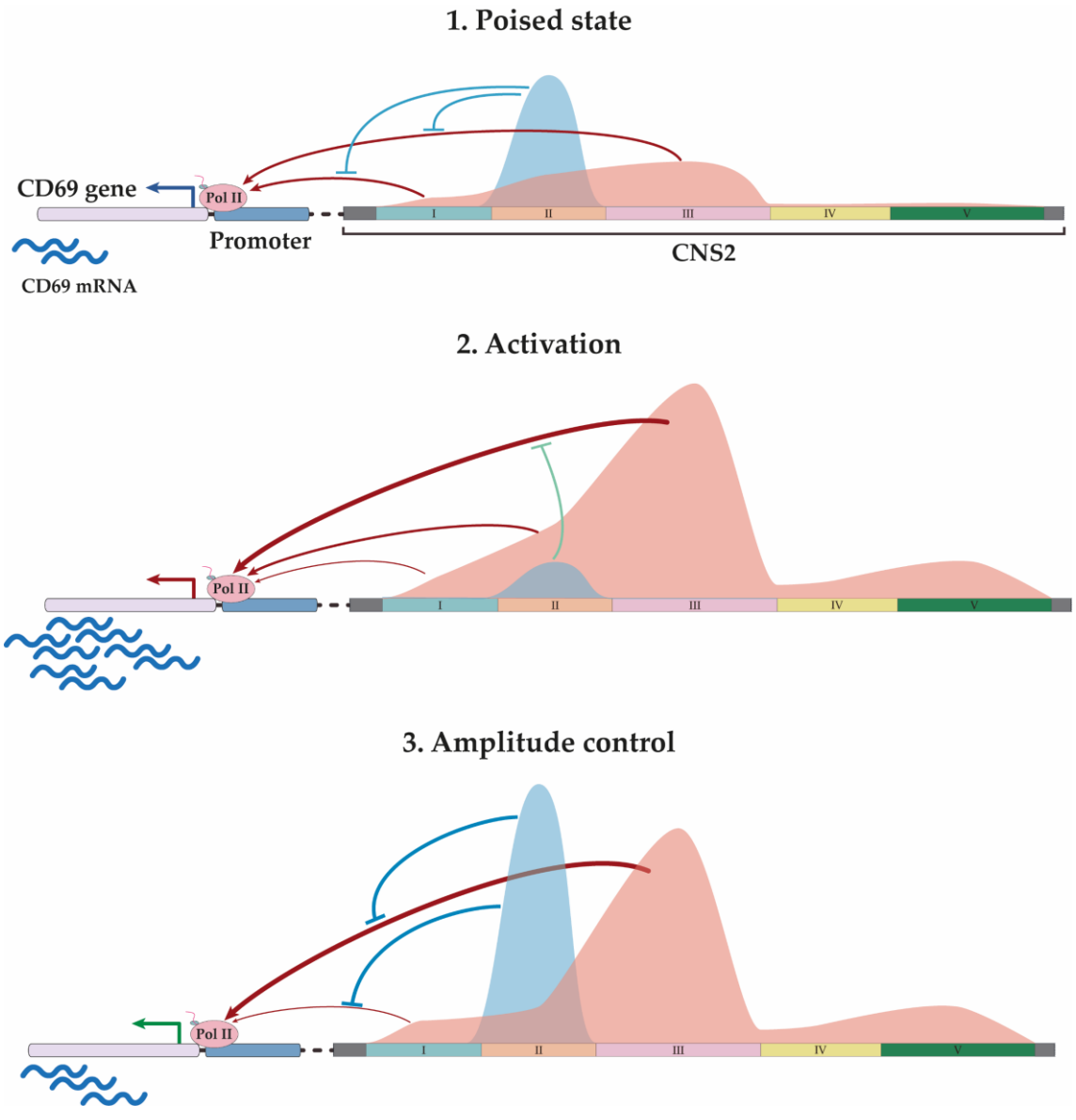


Figure I.15: CNS2 enhancer acts as a bivalent regulator of *CD69* expression. A three-step mechanism for CNS2 enhancer to turn on CD69 expression is proposed: **1)** a 'poised state' step, where the *locus* becomes poised for expression as repressive cofactors are recruited to the Core E, **2)** an 'activation' step, where the *locus* became quickly activated upon stimulation due to the recruitment of activation TFs both to the Core E and the 5'C region, and **3)** an 'amplitude control' step to modulate *CD69* gene expression mediated by the Chd4 recruitment to the Core E.

The presented data on the enrichment of Chd4 helicase suggest that it may have a prominent role in the Core E regulatory role since, in addition to its classical described function as part of the NuRD chromatin remodeling complex as a classical repressor **[144, 145]**, there has been accumulating evidence suggesting its NuRD-independent role as an activator of transcription **[115, 116]**. Moreover, it has been described a versatile use of NuRD by lineage-specific transcription factors to activate and repress gene transcription in the appropriate cellular and genetic context **[126, 144]**. As Oct1-NuRD may function as a repressive complex for the *IL-2* gene expression in mouse T cells **[129]**, we hypothesize that this complex may act similarly regulating the Core E activity, which is supported by the ENCODE data showing the enrichment of other NuRD components such as MTA3 within the Core E sequence. However, the binding dynamics of these TFs at shorter times need to be further investigated to address their repressive or activating effect on *CD69* transcription.

Also, in addition to exhibit a high expression of *CD69* at steady state and after stimulation, we have shown that the lack of the Core E in T cells led to an upregulation of the chemokine receptors CCR1, CCR2 and CCR5 at steady state, whereas upon stimulation their ligands CCL3L3, CCL4 and CCL4L1 were upregulated. These observations are in accordance with those of previous studies describing that CD69$^{-/-}$ mice show a dysregulation of chemokine and chemokine receptors production **[38, 39, 54, 146]**, and that mice with induced CD69 overexpression or treated with α-CD69 mAbs show an altered expression of pro inflammatory cytokines **[37, 147]**. Thus, *CD69* may be involved in the signaling pathway of chemokines and their receptors. Since human *CCR1*, *CCR2* and *CCR5* are located in the same gene cluster at chromosome 3, and *CCL3L3*, *CCL4* and *CCL4L1* are located in the same gene cluster at chromosome 17, we cannot rule out an inter-chromosomal interaction between the Core E and the abovementioned gene *loci* or with another genomic location which interacts with those chemokine clusters.

Overall, this work creates a comprehensive functional map of the TFs that bind to mouse and human CNS2, thus paving the way for understanding the fine-tuned CD69 expression regulation by the CNS2 regulatory Core in T cells.

Supplementary Figures I

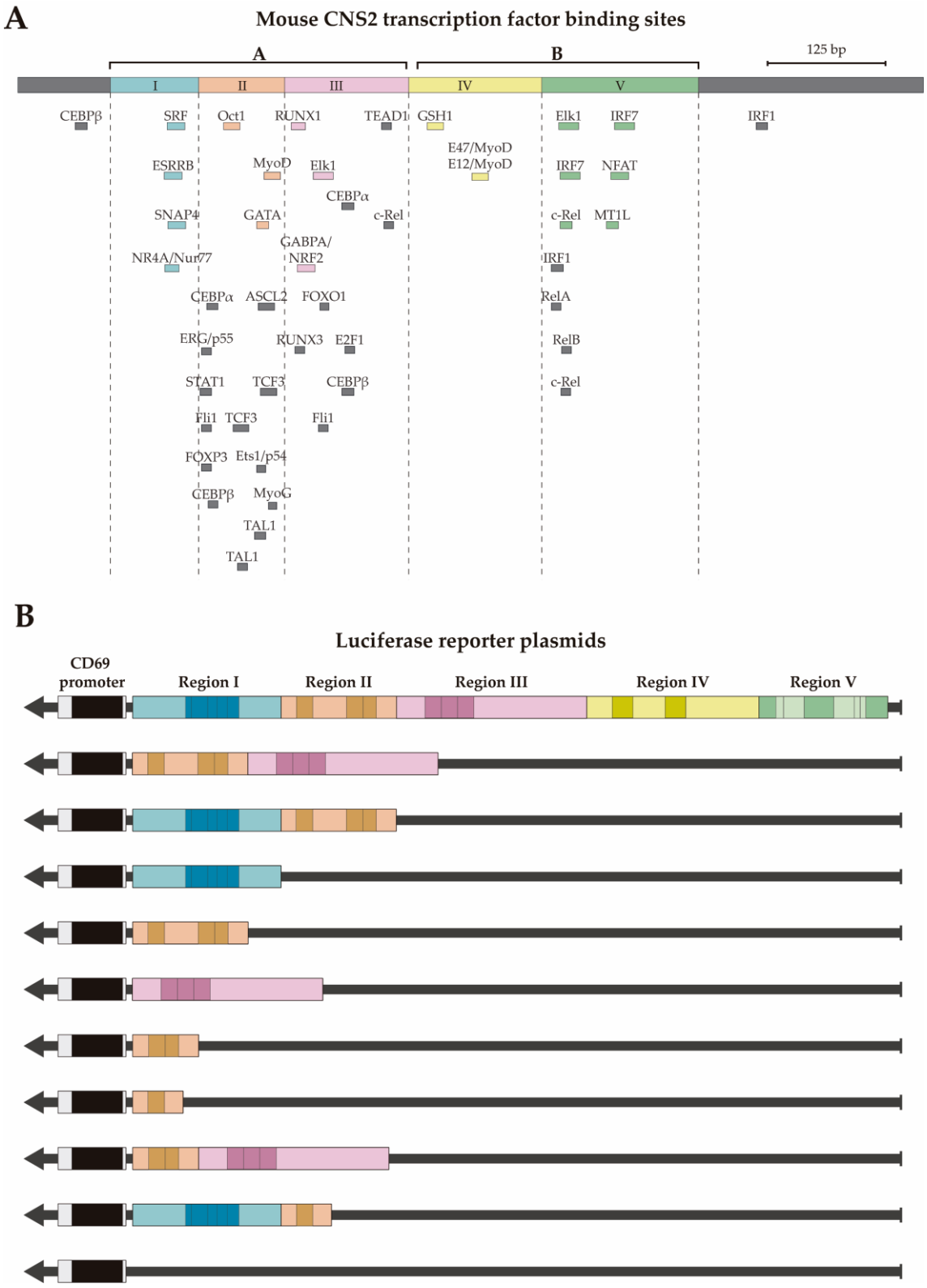


Figure SI.1: Luciferase reporter plasmids. A) Map of the CNS2 regions and TFBSs defined by available ChIP-Seq data. Color boxes indicate the TFBSs conserved among six mammal species; grey boxes indicate mouse-specific TFBSs. **B)** Luciferase reporter plasmids used in the reporter assay experiments.

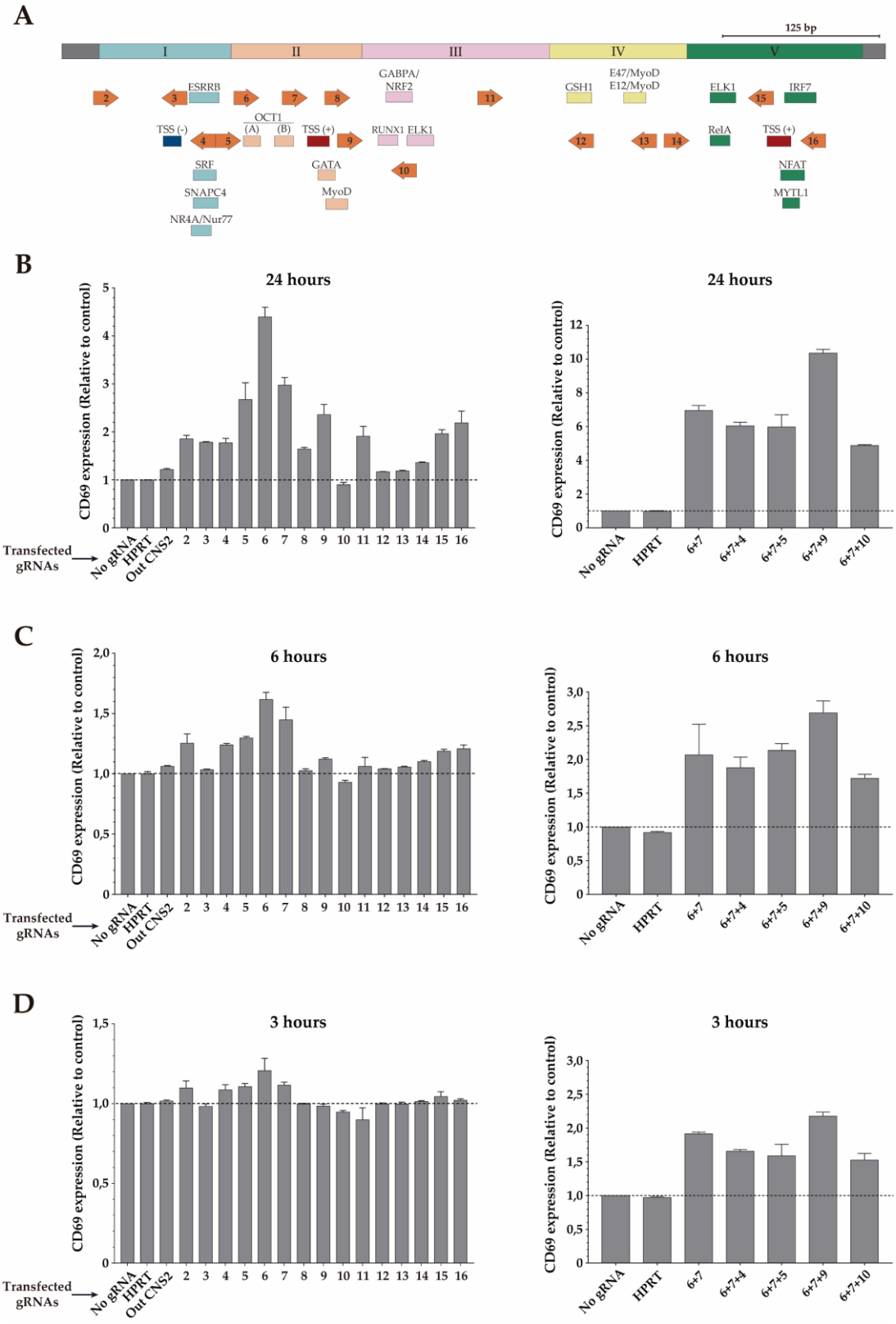


Figure SI.2: Transcriptional activation of CD69 expression. A) Schematic representation of CNS2 and the gRNAs used to recruit the dCas9-VP64 transcriptional activator. **B-D)** CD69 protein expression 24 hours (**B**), 6 hours (**C**) and 3 hours (**D**) after gRNA transfection of single (left panel) or combined (right panel) gRNAs targeting the main TFBSs within CNS2 in Jurkat CRISPR-a T cells. Data are presented as a mean ± SD of at least three independent experiments for each deleted region.

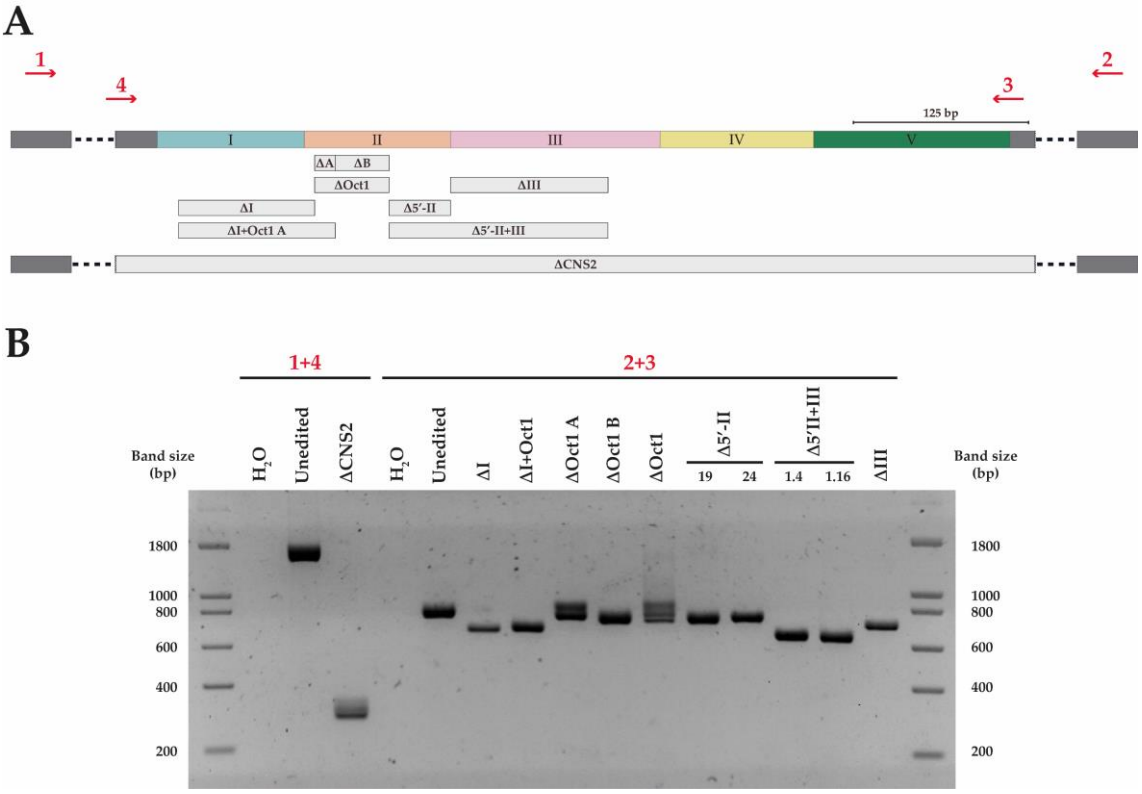


Figure SI.3: PCR screening of Jurkat T cells after CRISPR/Cas9 genome editing. A) Schematic representation of the CNS2 regions and the PCR primers used for the screening of the genome-edited Jurkat T cells. Numbers indicate the primers described in **Table I.7.** **B)** Representative agarose electrophoresis of the expected amplicons.

```
UneditedCNS2         1 ATAATATGTAACTTTCCATTTTAGGGAGACATTTTATGTGTCTTCTTTCATAAATGATAT
DelregionI           1 ATAATATGTAACTTTCCATTTTAGGGAGACATTTTATGTGTCTTCTTTCATAAATGATAT
DelregionI+Oct1      1 ATAATATGTAACTTTCCATTTTAGGGAGACATTTTATGTGTCTTCTTTCATAAATGATAT
DelOct1A             1 ATAATATGTAACTTTCCATTTTAGGGAGACATTTTATGTGTCTTCTTTCATAAATGATAT
DelOct1B             1 ATAATATGTAACTTTCCATTTTAGGGAGACATTTTATGTGTCTTCTTTCATAAATGATAT
DelCore E            1 ATAATATGTAACTTTCCATTTTAGGGAGACATTTTATGTGTCTTCTTTCATAAATGATAT
DelGATA+MyoD         1 ATAATATGTAACTTTCCATTTTAGGGAGACATTTTATGTGTCTTCTTTCATAAATGATAT
DelRUNX1+GATA+M      1 ATAATATGTAACTTTCCATTTTAGGGAGACATTTTATGTGTCTTCTTTCATAAATGATAT
DelregionIII         1 ATAATATGTAACTTTCCATTTTAGGGAGACATTTTATGTGTCTTCTTTCATAAATGATAT

UneditedCNS2        61 GTTAAAAAATTAAGTGATATTTTTAGGGCAGAAATTATCATTCACATTTCATCGTGTAGC
DelregionI          61 GTTAAAAAATTAAGTGATATTTTTAGGGCAGAAATTATCATTCACATTTCATCGTGTAGC
DelregionI+Oct1     61 GTTAAAAAATTAAGTGATATTTTTAGGGCAGAAATTATCATTCACATTTCATCGTGTAGC
DelOct1A            61 GTTAAAAAATTAAGTGATATTTTTAGGGCAGAAATTATCATTCACATTTCATCGTGTAGC
DelOct1B            61 GTTAAAAAATTAAGTGATATTTTTAGGGCAGAAATTATCATTCACATTTCATCGTGTAGC
DelCore E           61 GTTAAAAAATTAAGTGATATTTTTAGGGCAGAAATTATCATTCACATTTCATCGTGTAGC
DelGATA+MyoD        61 GTTAAAAAATTAAGTGATATTTTTAGGGCAGAAATTATCATTCACATTTCATCGTGTAGC
DelRUNX1+GATA+M     61 GTTAAAAAATTAAGTGATATTTTTAGGGCAGAAATTATCATTCACATTTCATCGTGTAGC
DelregionIII        61 GTTAAAAAATTAAGTGATATTTTTAGGGCAGAAATTATCATTCACATTTCATCGTGTAGC

UneditedCNS2       121 TGAGAAAATAGTATAGCCGAGGCTTAGAGAAATTAGTGCGGGGACAAGGTTGGCAAAAGT
DelregionI         121 TGAGAAAATAGTATAGCCGAGGCTTAGAGAAATTAGTGCGA------------------
DelregionI+Oct1    121 TGAGAAAATAGTATAGCCGAA--------------------------------------
DelOct1A           121 TGAGAAAATAGTATAGCCGAGGCTTAGAGAAATTAGTGCGGGGACAAGGTTGGCAAAAGT
DelOct1B           121 TGAGAAAATAGTATAGCCGAGGCTTAGAGAAATTAGTGCGGGGACAAGGTTGGCAAAAGT
DelCore E          121 TGAGAAAATAGTATAGCCGAGGCTTAGAGAAATTAGTGCGGGGACAAGGTTGGCAAAAGT
DelGATA+MyoD       121 TGAGAAAATAGTATAGCCGAGGCTTAGAGAAATTAGTGCGGGGACAAGGTTGGCAAAAGT
DelRUNX1+GATA+M    121 TGAGAAAATAGTATAGCCGAGGCTTAGAGAAATTAGTGCGGGGACAAGGTTGGCAAAAGT
DelregionIII       121 TGAGAAAATAGTATAGCCGAGGCTTAGAGAAATTAGTGCGGGGACAAGGTTGGCAAAAGT

UneditedCNS2       181 CTTTCAGGACACCAAAACTCAATGCAGCTTTCCTTCCTGTGTGACCTTATATGTAGAATC
DelregionI         162 ------------------------------------------------------------
DelregionI+Oct1    142 ------------------------------------------------------------
DelOct1A           181 CTTTCAGGACACCAAAACTCAATGCAGCTTTCCTTCCTGTGTGACCTTATATGTAGAATC
DelOct1B           181 CTTTCAGGACACCAAAACTCAATGCAGCTTTCCTTCCTGTGTGACCTTATATGTAGAATC
DelCore E          181 CTTTCAGGACACCAAAACTCAATGCAGCTTTCCTTCCTGTGTGACCTTATATGTAGAATC
DelGATA+MyoD       181 CTTTCAGGACACCAAAACTCAATGCAGCTTTCCTTCCTGTGTGACCTTATATGTAGAATC
DelRUNX1+GATA+M    181 CTTTCAGGACACCAAAACTCAATGCAGCTTTCCTTCCTGTGTGACCTTATATGTAGAATC
DelregionIII       181 CTTTCAGGACACCAAAACTCAATGCAGCTTTCCTTCCTGTGTGACCTTATATGTAGAATC

UneditedCNS2       241 CAGGGTGAGACGTCAGAAAGGAAGTAAAGATGAGGTAATATGAGCTAGATGCAAAGTGTG
DelregionI         162 ------------------GGAAGTAAAGATGAGGTAATATGAGCTAGATGCAAAGTGTG
DelregionI+Oct1    142 ----------------------------------GGTAATATGAGCTAGATGCAAAGTGTG
DelOct1A           241 CAGGGTGAGACGTCAGAAA--------------GGTAATATGAGCTAGATGCAAAGTGTG
DelOct1B           241 CAGGGTGAGACGTCAGAAAGGAAGTAAAGAT-----------------------------
DelCore E          241 CAGGGTGAGACG------------------------------------------------
DelGATA+MyoD       241 CAGGGTGAGACGTCAGAAAGGAAGTAAAGATGAGGTAATATGAGCTAGATGCAAAGTGTG
DelRUNX1+GATA+M    241 CAGGGTGAGACGTCAGAAAGGAAGTAAAGATGAGGTAATATGAGCTAGATGCAAAGTGTG
DelregionIII       241 CAGGGTGAGACGTCAGAAAGGAAGTAAAGATGAGGTAATATGAGCTAGATGCAAAGTGTG

UneditedCNS2       301 CAGAGAAGGTGGGGTACTCTGTGCTTATCAGACAGCTGCAGCAGCAGGATGCTTTTGGGT
DelregionI         203 CAGAGAAGGTGGGGTACTCTGTGCTTATCAGACAGCTGCAGCAGCAGGATGCTTTTGGGT
DelregionI+Oct1    169 CAGAGAAGGTGGGGTACTCTGTGCTTATCAGACAGCTGCAGCAGCAGGATGCTTTTGGGT
DelOct1A           287 CAGAGAAGGTGGGGTACTCTGTGCTTATCAGACAGCTGCAGCAGCAGGATGCTTTTGGGT
DelOct1B           272 ----------GGGGTACTCTGTGCTTATCAGACAGCTGCAGCAGCAGGATGCTTTTGGGT
DelCore E          253 -----------GGGTACTCTGTGCTTATCAGACAGCTGCAGCAGCAGGATGCTTTTGGGT
```

```
UneditedCNS2      361 ACATGTGAACAGGGACCACAGACTTCCGACCTAAGACAATTTCCGGAGACAAACTGCTCA
DelregionI        263 ACATGTGAACAGGGACCACAGACTTCCGACCTAAGACAATTTCCGGAGACAAACTGCTCA
DelregionI+Oct1   229 ACATGTGAACAGGGACCACAGACTTCCGACCTAAGACAATTTCCGGAGACAAACTGCTCA
DelOct1A          347 ACATGTGAACAGGGACCACAGACTTCCGACCTAAGACAATTTCCGGAGACAAACTGCTCA
DelOct1B          322 ACATGTGAACAGGGACCACAGACTTCCGACCTAAGACAATTTCCGGAGACAAACTGCTCA
DelCore E         302 ACATGTGAACAGGGACCACAGACTTCCGACCTAAGACAATTTCCGGAGACAAACTGCTCA
DelGATA+MyoD      317 ACATGTGAACAGGGACCACAGACTTCCGACCTAAGACAATTTCCGGAGACAAACTGCTCA
DelRUNX1+GATA+M   301 ------------------------------------------------------------
DelregionIII      357 ------------------------------------------------------------

UneditedCNS2      421 TCTCTCTTTTTTTTTTTTTCCTTTCAGTGGGTGAATTAGGTTTCTGAATGGCTTTTAATC
DelregionI        323 TCTCTCTTTTTTTTTTTTTCCTTTCAGTGGGTGAATTAGGTTTCTGAATGGCTTTTAATC
DelregionI+Oct1   289 TCTCTCTTTTTTTTTTTTTCCTTTCAGTGGGTGAATTAGGTTTCTGAATGGCTTTTAATC
DelOct1A          407 TCTCTCTTTTTTTTTTTTTCCTTTCAGTGGGTGAATTAGGTTTCTGAATGGCTTTTAATC
DelOct1B          382 TCTCTCTTTTTTTTTTTTTCCTTTCAGTGGGTGAATTAGGTTTCTGAATGGCTTTTAATC
DelCore E         362 TCTCTCTTTTTTTTTTTTTCCTTTCAGTGGGTGAATTAGGTTTCTGAATGGCTTTTAATC
DelGATA+MyoD      377 TCTCTCTTTTTTTTTTTTTCCTTTCAGTGGGTGAATTAGGTTTCTGAATGGCTTTTAATC
DelRUNX1+GATA+M   301 ----------------------------------------------------TGGCTTTTAATC
DelregionIII      357 ----------------------------------------------------GGCTTTTAATC

UneditedCNS2      481 TTCTTGTGAATCAAAACTTCAATTTAGAATACATCATCCTCTCTCAAATTAATGCAAATT
DelregionI        383 TTCTTGTGAATCAAAACTTCAATTTAGAATACATCATCCTCTCTCAAATTAATGCAAATT
DelregionI+Oct1   349 TTCTTGTGAATCAAAACTTCAATTTAGAATACATCATCCTCTCTCAAATTAATGCAAATT
DelOct1A          467 TTCTTGTGAATCAAAACTTCAATTTAGAATACATCATCCTCTCTCAAATTAATGCAAATT
DelOct1B          442 TTCTTGTGAATCAAAACTTCAATTTAGAATACATCATCCTCTCTCAAATTAATGCAAATT
DelCore E         422 TTCTTGTGAATCAAAACTTCAATTTAGAATACATCATCCTCTCTCAAATTAATGCAAATT
DelGATA+MyoD      437 TTCTTGTGAATCAAAACTTCAATTTAGAATACATCATCCTCTCTCAAATTAATGCAAATT
DelRUNX1+GATA+M   313 TTCTTGTGAATCAAAACTTCAATTTAGAATACATCATCCTCTCTCAAATTAATGCAAATT
DelregionIII      368 TTCTTGTGAATCAAAACTTCAATTTAGAATACATCATCCTCTCTCAAATTAATGCAAATT

UneditedCNS2      541 ACACAAATTGGTAAAATTTATTTTTAACCTGCAGGTGCCTTTGTGTTTTTGAATTGTTTA
DelregionI        443 ACACAAATTGGTAAAATTTATTTTTAACCTGCAGGTGCCTTTGTGTTTTTGAATTGTTTA
DelregionI+Oct1   409 ACACAAATTGGTAAAATTTATTTTTAACCTGCAGGTGCCTTTGTGTTTTTGAATTGTTTA
DelOct1A          527 ACACAAATTGGTAAAATTTATTTTTAACCTGCAGGTGCCTTTGTGTTTTTGAATTGTTTA
DelOct1B          502 ACACAAATTGGTAAAATTTATTTTTAACCTGCAGGTGCCTTTGTGTTTTTGAATTGTTTA
DelCore E         482 ACACAAATTGGTAAAATTTATTTTTAACCTGCAGGTGCCTTTGTGTTTTTGAATTGTTTA
DelGATA+MyoD      497 ACACAAATTGGTAAAATTTATTTTTAACCTGCAGGTGCCTTTGTGTTTTTGAATTGTTTA
DelRUNX1+GATA+M   373 ACACAAATTGGTAAAATTTATTTTTAACCTGCAGGTGCCTTTGTGTTTTTGAATTGTTTA
DelregionIII      428 ACACAAATTGGTAAAATTTATTTTTAACCTGCAGGTGCCTTTGTGTTTTTGAATTGTTTA

UneditedCNS2      601 CCGTAAAGGGTCAAT
DelregionI        503 CCGTAAAGGGTCAAT
DelregionI+Oct1   469 CCGTAAAGGGTCAAT
DelOct1A          587 CCGTAAAGGGTCAAT
DelOct1B          562 CCGTAAAGGGTCAAT
DelCore E         542 CCGTAAAGGGTCAAT
DelGATA+MyoD      557 CCGTAAAGGGTCAAT
DelRUNX1+GATA+M   433 CCGTAAAGGGTCAAT
DelregionIII      488 CCGTAAAGGGTCAAT
```

Figure SI.4: Multiple alignment of selected CRISPR/Cas9 edited Jurkat cell lines. After gene editing, Jurkat cell lines were analyzed by PCR and PCR amplicons were sequenced by Sanger sequencing. A selection of cells carrying the desired deletions are shown.

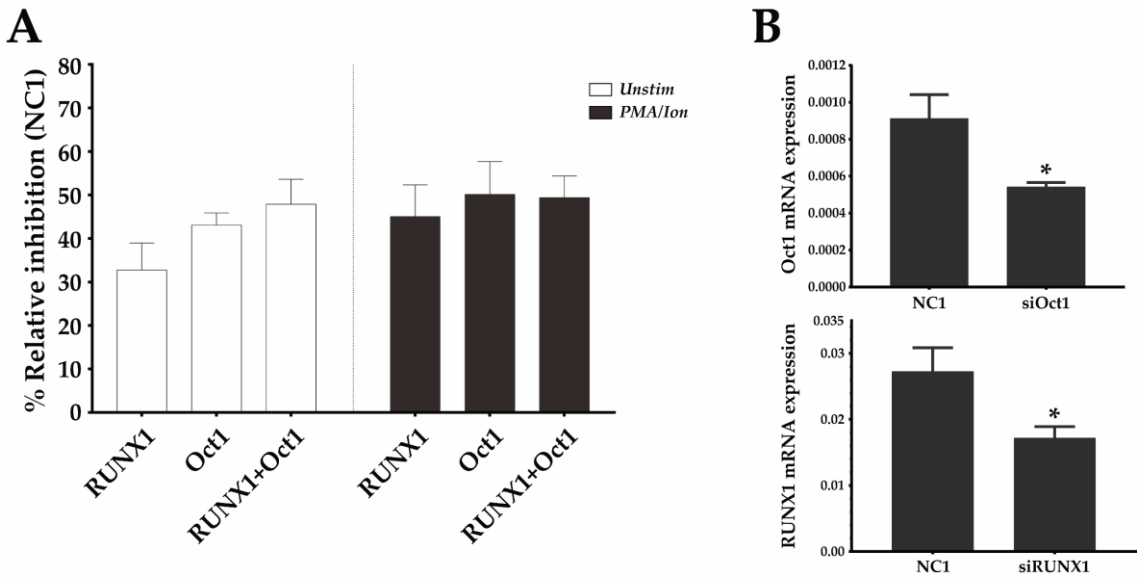


Figure SI.5: *RUNX1* and *Oct1* silencing downregulates mouse CD69 protein expression. **(A)** Mouse EL-4 T cells were transfected with a 1 μM pool of three DsiRNAs against *RUNX1*, *Oct1*, a combination of both, or a non-targeting DsiRNA (NC-1) and cultured for 16 hours. They were then stimulated or not with PMA/Io, and CD69 expression was measured 6h later. The bars represent the mean ± SEM of the % relative inhibition with respect to CD69 expression in the NC-1 control of 2 different experiments in which each transfection was performed in duplicate. **(B)** mRNA expression of *RUNX1* and *Oct1* was analyzed by qPCR 22 hours after DsiRNA transfection, compared with the NC1 DsiRNA control. Data are presented as mean ± SD of two independent experiments (*P < 0.05).

Chapter II:

Analysis of Antigen Presentation by Genome-Edited *ERAP2*$^{-/-}$ B Lymphocytes

Introduction II

1. Antigen Processing and Peptide Presentation by MHC Class I Molecules

Major histocompatibility class I molecules (MHC class I, HLA in humans) are trimeric complexes made of by a heavy (HC) and a light chain ($\beta_2$microglobulin) and an antigenic peptide. MHC class I molecules present peptide antigens to $CD8^+$ T lymphocytes. Self and pathogenic proteins are proteolytically degraded by proteasomes and other peptidases in the cytosol to generate a pool of peptides, which is extremely diverse both in sequence and length. The transporter associated with antigen processing (TAP) specifically translocate some of these degradation products to the endoplasmic reticulum lumen [148]. Peptides with the correct length (8-11 residues) interact with specific pockets of the antigen recognition site of the MHC class I molecule [149]. Usually position 2 (hereafter P2) is the main anchor residue and/or the C-terminus residue (PΩ) or other positions of the peptide [150, 151]. These interactions help to optimize the stability of the class I complex. The fraction of amino terminally extended precursors can also be used for antigen presentation after precursor editing and trimming by endoplasmic reticulum (ER) resident aminopeptidase activities.

2. Human Endoplasmic Reticulum Aminopeptidases

In humans, two related endoplasmic reticulum aminopeptidases called ERAP1 [152] and ERAP2 [153, 154] with non-redundant specificities and different substrate preferences trim N-terminally extended precursors of these peptides to their final length. ERAP1 has a wider specificity since it cleaves virtually all N-terminal residues [155] except proline (Pro), but shows a wide range of efficiency depending on the N-terminal side chain of the substrate. ERAP1 shows preference for nonpolar residues, being much less efficient with polar and charged ones [155], and preferentially cleaves peptides longer than 9-mers [156]. In contrast, ERAP2 activity preferentially targets basic residues, being more efficient cleaving arginine (Arg), and cleaves best 9-mers and shorter peptides [154, 156-160]. Peptide binding to HLA class I molecules in the ER stabilizes the nascent HC/$\beta_2$microglobulin complexes and allows their subsequent transport to the cell surface, where they are recognized by the $CD8^+$ cytotoxic T lymphocytes (CTLs) [161]. The recognition of foreign or self-peptide ligands can lead to the killing of pathogen infected cells or, instead, to initiate an autoimmune damage, respectively.

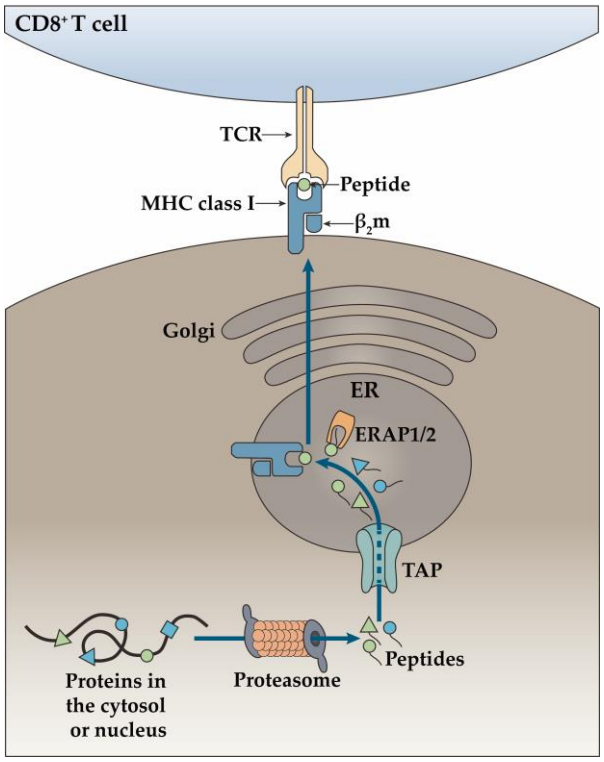


Figure II.1: Basic MHC class I antigen presentation pathway. Presentation of intracellular antigenic peptides by MHC class I molecules occurs as a result of a series of processes. Peptides resulting from the protein degradation by the proteasome are translocated via TAP into the ER lumen. Then, peptides are process by ERAP1/2 aminopeptidases and loaded onto MHC class I molecules. Peptide-MHC class I complexes are released from the ER and transported to the plasma membrane for antigen presentation to $CD8^+$ T cells. Adapted from [162].

3. HLA-B*27 Class I Allele and Its Relationship with Autoimmunity

HLA-B*27 class I allele prevalence varies significantly among different subpopulations, being more common in Caucasoids (Allele Frequency Net Database [163], available at allelefrequencies.net). HLA-B*27 is strongly associated and plays a central role in the pathogenesis of many spondyloarthropaties, in particular in the ankylosing spondylitis (AS) [164], a chronic arthritis that primarily affects the spinal and sacroiliac joints, although other joints can also be involved [165]. There is a differential association of HLA-B*27 subtypes to AS. B*27:05, B*27:02 and B*27:04 are strongly associated to this disease. B*27:07 is also associated to AS in multiple, but not all, populations, and B*27:06 and B*27:09 are not or very weakly associated to AS [166]. These subtypes differ in a few amino acid positions located in the peptide binding groove.

Several hypotheses have been put forward to elucidate the mechanism by which HLA-B*27 contributes to the pathogenesis of spondyloarthritis, many of them based on the idea that the classical function of HLA-B*27 as an antigen presenting molecule leads to the development of inflammatory diseases through molecular mimicry with bacterial antigens. However, the bases for this association remain largely unknown, and none of these hypotheses has reasonably explained the mechanism and the differential association between HLA-B*27 subtypes and the disease association.

The classical arthrogenic peptide hypothesis [167] is based on the antigen presenting properties of HLA-B*27. During bacterial infection, microbial peptide epitopes bound to HLA-B*27 elicit a normal CTL response against the pathogen. This hypothesis assumes that some of these antimicrobial effector T cells would cross-react with autologous self-peptide ligands, which are also presented by the HLA-B*27 class I molecule, showing molecular mimicry with the bacterial epitope. Thus, the undesired cross-reaction would trigger autoimmune tissue injury and inflammation.

Under this assumption, variations in HLA binding or immune recognition between self-peptides and bacterial epitopes with the HLA-B*27 subtypes would explain their differential association with AS. In recent years both ERAP1 and ERAP2 have been associated with several inflammatory diseases. ERAP1 has been associated with AS, psoriasis and Behçet's disease in epistasis with the risk HLA-B*27 [168], B*40:01 [169], C*06:02 [170] and B*51 alleles [171], respectively. ERAP2 is associated with AS and psoriasis [172-175], but not in epistasis with MHC class I. Both ERAP1 and ERAP2 are also risk factors for birdshot chorioretinopathy [176, 177]. As ERAP1 and ERAP2 have been described as one of the main non HLA susceptibility genes for AS [178-180], this has led to renewed interest in the arthrogenic peptide hypothesis.

Two previous studies using unrelated ERAP2$^{+/+}$ and ERAP2$^{-/-}$ lymphoblastoid cell lines (LCLs) tried to determine the effect of ERAP2 aminopeptidase on the generation of HLA-B*27 peptide ligands [181, 182]. Moreover, although a rigorous pairwise comparison between both types of cell lines was carried out, these LCLs may vary in the expression of many proteins, including those involved in antigen processing, so the possibility of background differences cannot be formally ruled out.

To elucidate the role of ERAP2 aminopeptidase in the generation of HLA-B*27 ligands, high throughput mass spectrometry analyses of peptides isolated from CRISPR/Cas9 edited ERAP2$^{-/-}$ cell clones were carried out. In the unedited (ERAP2$^{+/+}$) and edited (ERAP2$^{-/-}$) isogenic cell clones, all the antigen presentation machinery and the rest of cellular proteins are common, except for the deleted ERAP2 protein. Results determined that ERAP2 selects HLA-B*27:05 ligands with a higher contribution of P2 anchor motifs, and lower influences of auxiliary residues.

Materials and Methods II

1. Cells and Antibodies

C1R-B27 is an HLA-B*27:05 transfectant [183] of the human lymphoid cell line HMy2.C1R. Cells were cultured in RPMI 1640 with 10% fetal bovine serum. Monoclonal antibodies specific for a monomorphic HLA class I (W6/32 [184]), for ERAP2 (R&D Systems, MAB3830), and for γ-tubulin (Sigma Aldrich, GTU88) were used.

2. Characterization of Human *ERAP2* Genomic *Locus*, crRNA Design and CRISPR/Cas9 RNP Transfection

Sequence and genomic *locus* distribution of human *ERAP2* were obtained from the UCSC Genome Browser [87] (GRCh38/hg38 database, ENCODE consortium [88]). Human ERAP2 (Chr5: 96,876,506 - 96,917,932) 5' and 3' UTRs, coding exons and introns were identified and depicted using Snapgene® (from GSL Biotech; available at snapgene.com).

Human *ERAP2* aminopeptidase exons 5 and 6 (chr5: 96,892,299 - 96,892,453) were analyzed using the CRISPR/Cas9 design tool from Benchling (Biology software, 2019. Retrieved from benchling.com). **Table II.1** shows the selected crRNAs, their sequence, targets and genomic location. Selection criteria and additional checks of the designed crRNAs were carried out as stated in **M&M I.5**. sgRNA generation and RNP transfection on C1R-B*27:05 cells were carried out as described in **M&M I.5**.

3. Deletion of Human *ERAP2* Exons 5-6 and Characterization of ERAP2 Expression

DNA of transfected C1R-B*27:05 cells was extracted 48 hours after RNP transfection, and *ERAP2 locus* was analyzed by PCR using primers that amplified inside and outside of the deleted region (**Table II.2** and **Figure II. 4**). Transfected cells were then cloned, and ERAP2 expression was analyzed by PCR, qPCR (**Table II.3** and **Figure II.5**) and Western Blot (**Figure II.5**) as described [185].

Two unedited (UN17 and UN20, with the complete *ERAP2* sequence) and two homozygously edited (KO9 and KO19, without *ERAP2* exons 5 and 6) cell clones were selected for further characterization.

4. HLA Class I-Bound Peptide Isolation

HLA-bound peptides were isolated from 10^9 UN17, UN20, KO9 and KO19 cell clones, as described [186]. Three independent biological replicates were obtained. Cells were lysed in 1% Igepal CA-630 (Sigma-Aldrich I8896), 20 mM Tris/HCl, and 150 mM NaCl (pH 7.5) in the presence of a protease inhibitor cocktail (Sigma-Aldrich, 04693116001). After centrifugation, supernatants were first passed through a control precolumn containing CNBr-activated sepharose 4B (GE Healthcare, Buckinghamshire, UK) to remove non-specific peptides. The HLA-B*27:05/peptide complexes were isolated by affinity chromatography from the soluble cell extract fraction with the anti-HLA class I W6/32 monoclonal antibody [184]. HLA-bound peptides were eluted with 1% aqueous trifluoroacetic acid (TFA) and concentrated by ultra-filtration with a vivaspin 2 filter (5,000 MWCO HY, Sartorius Stedim Biotech, Goettingen, Germany), as previously described [187].

5. Electrospray-Ion Trap Mass Spectrometry Analysis

Peptide mixtures were desalted using OMIX Tips (C_{18}, Agilent Technologies, Santa Clara, CA) and analyzed by nanoLC-MS/MS using a Q-Exactive-Plus mass spectrometer fitted with an Ultimate 3000 RSLC nanocapillary UHPLC (Thermo Fisher Scientific) using the same parameters as previously described [188]. Peptides were resolved on a Reprosil C18-Aqua capillary columns (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany) with a 5-28% acetonitrile linear gradient for 2 hours in presence of 0.1% formic acid at a flow rate of 0.15 µl/minute. The dynamic exclusion was set to 20 seconds and the automatic gain control value for the full MS was set to 3×10^6. The selected masses were fragmented from the survey scan of mass-to-charge ratio (m/z) 300-1,800 AMU at resolution 70,000. The 10 most intense masses from each full mass spectrum were fragmented by higher-energy collisional dissociation. MS/MS spectra were acquired with a resolution of 17,500 at m/z 200. The target values of the MS/MS were set to 1×10^5 and the isolation window to 1.8 m/z. The maximum injection time was set to 100 ms and normalized collision energy to 25 eV.

6. Experimental Design

The effect of ERAP2 depletion on the HLA-B*27:05 peptidome was analyzed by quantitative label-free mass spectrometry of two unedited (with the complete ERAP2 sequence) and two homozygously edited (without *ERAP2* exons 5 and 6, which encodes the active site of the protein) cell clones. Three independent preparations of the HLA-B*27:05 bound peptides were obtained from each one of the four cell lines analyzed, and biological replicates were performed. The correlation in peptide intensities between UN and KO cell clones was evaluated by Pearson correlation analyses (**Figure II.2**).

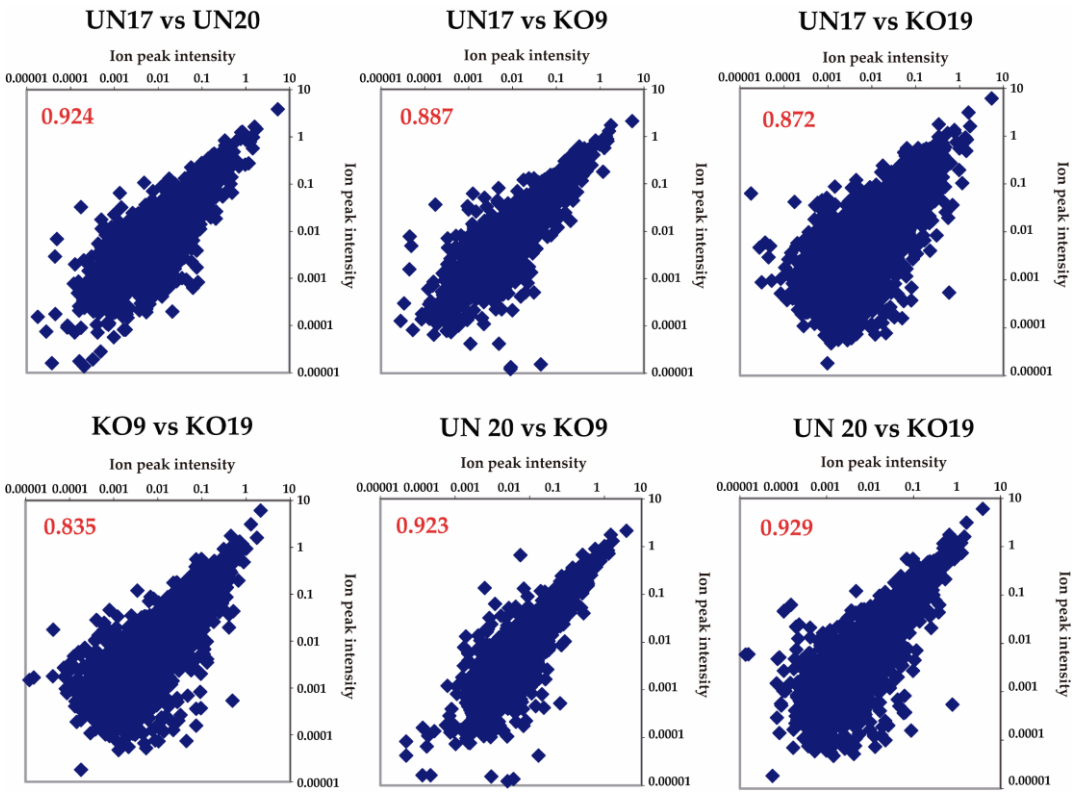


Figure II.2: Pearson correlation analyses of identified peptides between UN and KO cell clones. Pairwise comparisons of the ion peaks intensities corresponding to the assigned peptides from UN and KO cell clones of HLA-B*27:05 ligands. Pearson correlation values are indicated in red.

A precolumn was used to remove the non-specific binding proteins and peptides, and a 0.01 false discovery rate (FDR) was used. Since a majority of the identified peptides were found in both the UN and KO cells (**Figure II.3**), quantitative differences in peptide amounts between cell lines were assigned in pairwise comparisons as follows **[189]**: in each experiment, the intensity of a given ion peak was normalized to the total intensity of all the identified B*27:05 ligands. The mean normalized intensity of each ion peak from

the three individual experiments for each cell line was taken as the amount of that peptide relative to the total amount of ligands identified in that cell clone. HLA-B*27:05 ligands in each pairwise comparison were classified based on the normalized intensity ratio (IR) of the corresponding ion peak in the cell clones. Peptides predominant in one cell clone, relative to the other (IR > 1), were subdivided into two subsets as previously described [189]. Peptides with IR > 1 to 1.5 in every cell line were considered to be expressed in similar amounts in the two cell clones, and therefore affected a little or not at all by ERAP2. Peptides with IR > 1.5 in one cell clone relative to the other, including peptides found only in that cell clone, were considered to be upregulated and, consequently, influenced by ERAP2 context.

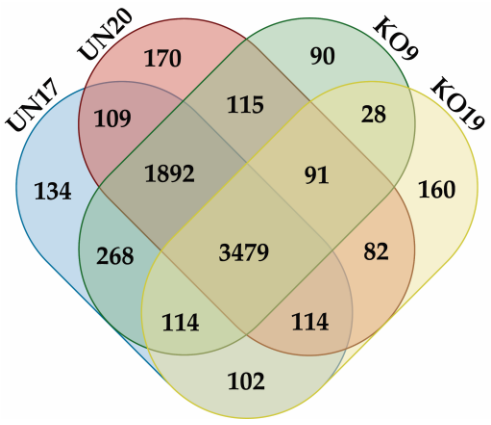


Figure II.3: Venn diagram for the identified HLA-B*27:05 ligands. Number of peptides identified in different UN and KO cell clones are indicated; overlapping sections show the peptides identified in more than one cell clone.

7. Peptide Database Searches and Statistical Analysis

Peptides were identified using MaxQuant software [190] (version 1.6.0.16, Max Planck Institute of Biotechnology) with the Andromeda [191] search engine using the human section of the UniProt/Swiss-Prot database (release 27.2.19, containing 73,101 entries). Methionine oxidation and N-acetylation were accepted as variable modifications. The peptide precursors and fragment mass tolerances were set at 6 and 20 ppm, respectively and the false discovery rate was 0.01. Peptides with the previously defined HLA*B27 motifs of Arg/Gln/Lys at P2 were selected if the MaxQuant score was greater than 45. Analysis of statistical significance of the differences in residue frequencies was performed by using a Bonferroni-corrected chi-square test. The relative intensity differences among peptide sets were analyzed with linear regression.

Multiple t-tests were used to analyze the differences in the hydrophobicity at different peptide positions among the HLA-B*27:05 ligands. P values <0.05 were considered to be statistically significant.

8. Tables

Table II.1: crRNAs designed for *ERAP2* genome editing, its sequences, targets and genomic location.

crRNA name	Sequence 5′ → 3′	PAM	Target	Genomic location, GRCh38/hg38
ERAP2_5′_exon5	CTTTGCTAATGATTTCACGT	AGG	*ERAP2* exon 5 5′	Chr5: 96,891,288 - 96,891,307
ERAP2_3′_exon6	ATACCATATCTCCCTTGCAG	AGG	*ERAP2* exon 6 3′	Chr5: 96,896,015 - 96,896,034

Table II.2: PCR primers used for *ERAP2* genotyping and the expected sizes of the amplicons.

Primer	Sequence 5′ → 3′	Amplicon length		Genomic location, GRCh38/hg38
		Unedited	ERAP2$^{-/-}$	
Ext_Fw	TGGCATACAGTTCCAGGCTT	5386 bp	639 bp	Chr5: 96,890,695 - 96,890,714
Ext_Rv	ACTTGTTACGTGCCTAGACCT			Chr5: 96,896,060 - 96,896,080
Int_Fw	TCCCTTTCCATGAATGCCCT	436 bp	No amplicon	Chr5: 96,894,296 - 96,894,315
Int_Rv	AACCTCTGCCCCTGACTTAC			Chr5: 96,894,702 - 96,894,721

Table II.3: SYBR green qPCR primers used for ERAP2 mRNA amplification.

Primer	Sequence 5′ → 3′	Targeting exons
β2M_Fw	GATGAGTATGCCTGCCGTGTG	*B2microglobulin*
β2M_Rv	CAATCCAAATGCGGCATCT	
ERAP2_Ex2_Fw	TCAAAGCAAGGGAAAGCCA	*ERAP2* exons
ERAP2_Ex3_Rv	AAGCACATACAGAACTCTTGGT	2-3
ERAP2_Ex3_Fw	CTTGAAGGAGGTCTTTTGGAAG	*ERAP2* exons
ERAP2_Ex4_Rv	GGGGATGCATAGATGGACAC	3-4
ERAP2_qPCR_Ex4_Fw	TCCATCTATGCATCCCCAGAC	*ERAP2* exons
ERAP2_qPCR_Ex5_Rv	TGGGCTATGACTCTGGTGAC	4-5
ERAP2_Ex5_Fw	CACTGCTTTTTGACCCCAAG	*ERAP2* exons
ERAP2_Ex6_Rv	CACCATTCCATTGTGACCAG	5-6
ERAP2_Ex6_Fw	CACAATGGAATGGTGGAATG	*ERAP2* exons
ERAP2_Ex7_Rv	TTCCGCTGGTTTGGAGATAG	6-7
ERAP2_Ex9_Fw	CTTTGACCTCTGCATTTTCCC	*ERAP2* exons
ERAP2_Ex11_Rv	CTATAGAAATGCTAAGAATGATGACTTGT	9-11
ERAP2_Ex14_Fw	TTCTCCCAGTGGATGGAATC	*ERAP2* exons
ERAP2_Ex15_Rv	GATGCTTGCTCGTTGACAAA	14-15

Results II

1. Human *ERAP2* Genomic *Locus* Characterization and *ERAP2* Functional KO in Human B Lymphocytes

In order to identify the target sequences for CRISPR/Cas9 genome editing of human *ERAP2* aminopeptidase, the UCSC Genome Browser with the human GRCh38/hg18 database was used. The human *ERAP2* gene is located adjacent to a closely related aminopeptidase gene (ERAP1) on chromosome 5, and its genomic sequence (chr5: 96,875,939 – 96,919,716) comprises ~44 kbp with 19 coding exons. As *ERAP2* exon 5 codifies for the enzyme active site and the exon 6 codes for a catalytic zinc binding domain, the easiest way to knock down ERAP2 function was to delete the exons 5 and 6. Deletion of exons 5 and 6 would cause a frameshift mutation leading to an early termination of the exon4/exon7 fusion (**Figure II.4 A**); moreover, if expressed, the truncated protein would not be enzymatically active.

The human B cell line C1R expressing the HLA-B*27:05 molecule was edited by using CRISPR/Cas9 technology to generate a functional KO of the ERAP2 aminopeptidase. *ERAP2* exons 5 and 6 were removed by deleting a 4741 bp fragment of the *ERAP2* gene (**Figure II.4 A**). Cells were cloned by limiting dilution and the correct editing of ERAP2 *locus* in C1R-B*27:05 cells was confirmed by PCR analysis (**Figure II.4 B**). Two UN and KO clones were selected for further experiments.

qPCR analysis of *ERAP2* mRNA in ERAP2$^{-/-}$ cells shown a mild reduction of transcript variants containing exons 3-4, a drastic decrease of transcript variants containing exons 2-3, 9-11 and 14-15, and a complete absence of transcript variants containing either exons 4-5, 5-6 or 6-7 (**Figure II.5 A**). These data support the hypothesis that deletion of exons 5 and 6 causes a frameshift mutation that leads to an early termination of ERAP2 protein. Western blot analysis of ERAP2 expression confirmed the absence of protein expression (**Figure II.5 B**). mRNA and protein expression analysis confirmed that ERAP2 aminopeptidase is absent in the edited clones, thus suggesting that it is feasible to efficiently delete long genomic regions with CRISPR/Cas9. The effect on surface expression of B*27:05 heterodimers was assessed by flow cytometry with the ME1 mAb, but statistically no differences in the KO cell clones, relative to the UN cell clones, were observed (data not shown).

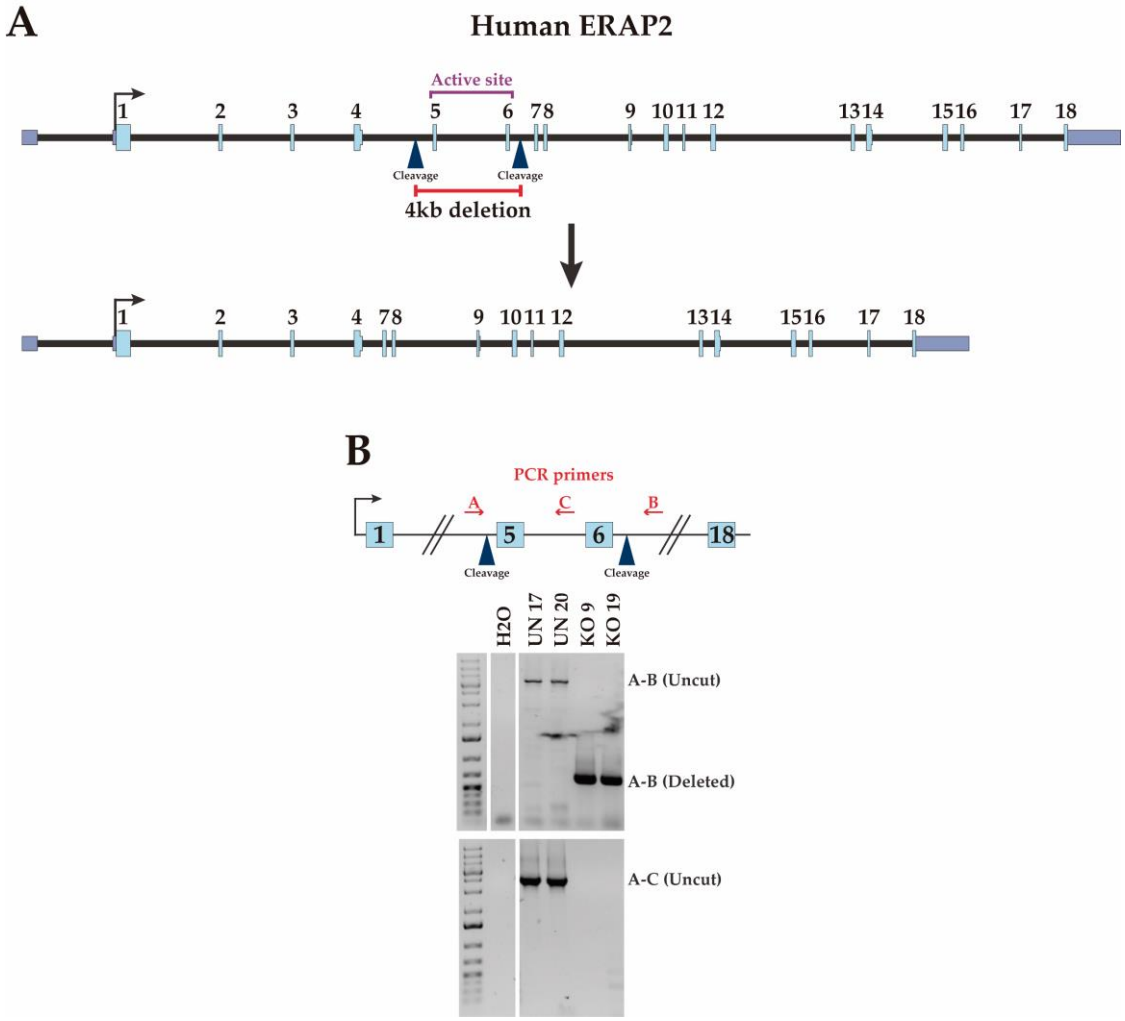


Figure II.4: Targeted deletion of *ERAP2* exons 5 and 6 in HLA-B*27:05 cells.
A) Schematic representation of *ERAP2* gene in the human chromosome 5 and CRISPR/Cas9 deletion of exons 5 and 6**. B)** Analysis of the deletion of *ERAP2* exons 5 and 6 by PCR from ERAP2$^{+/+}$ (UN) and ERAP2$^{-/-}$ (KO) cell clones using external (A-B) and internal (C) PCR primers.

2. HLA-B*27:05 Peptidome in ERAP2$^{+/+}$ and ERAP2$^{-/-}$ B Cells, and Effect of *ERAP2* in HLA-B*27:05 Peptidome

Peptide pools bound to HLA-B*27:05 class I molecules isolated from two UN (UN17 and UN20) and two KO (KO9 and KO19) cell clones were subjected to LCMS/MS analysis in biologically triplicate experiments. A total of 6,212 and 6,052 peptides were identified in UN17 and UN20 cell clones, respectively, and 6,077 and 4,170 peptides were identified in KO9 and KO19 cell clones, respectively. The anchor motifs for HLA-B*27:05 binding (Arg, Gln or Lys) at position 2 (P2) were detected in peptides from both UN and KO cell clones. *In silico* binding prediction differences between the HLA-B*27:05 ligands from UN and KO were not found (**Figure SII.1**).

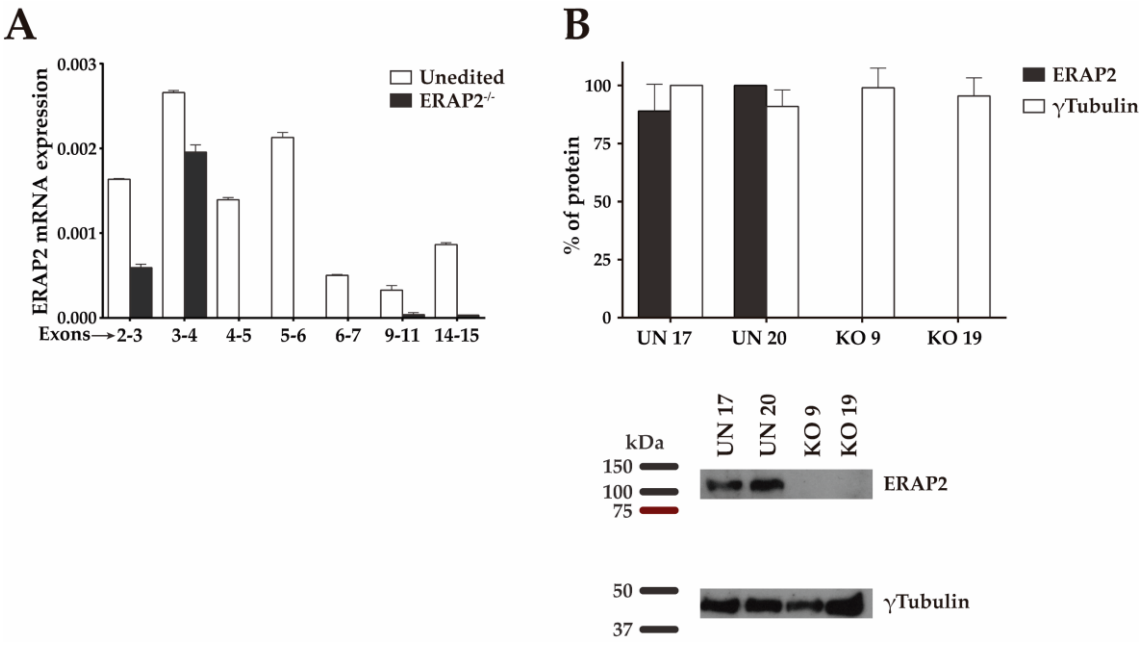


Figure II.5: *ERAP2* mRNA and protein expression in unedited (UN) and edited (KO) B cells.
A) *ERAP2* mRNA expression was measured with qPCR primers for exons 2-3, 3-4, 4-5, 5-6, 6-7, 9-11 and 14-15. Data are presented as mean ± SD of at least three independent experiments. **B)** Representative western blot and expression levels of ERAP2 (closed bars) and γ-tubulin (open bars) in the indicated UN and KO cell clones relative to the parental cell line showing maximal values.

To analyze the influence of ERAP2 on the relative amounts of HLA-B*27:05 ligands, comparison between ERAP2$^{+/+}$ and ERAP2$^{-/-}$ HLA-B*27:05 peptidomes was carried out. HLA-B*27:05 bound peptides were classified based on the relative intensity of the UN versus the KO in each of the two types of cell lines compared. Enriched peptides in presence or absence of *ERAP2* (with a normalized intensity ratio (IR) > 1.0 relative to the other condition) were compared and split into two subgroups: **a)** 1.0 < IR ≤ 1.5 (mild enrichment) and **b)** IR > 1.5 (strong enrichment), as the statistical significance of IR values was previously established [192]. Also, specific peptides only identified in UN or KO clones were included in their respective IR > 1.5 subgroup. No differences in MW were found when 1.0 < IR ≤ 1.5 from UN and KO were compared (Figure II.6, left panel). In contrast, the subset of HLA-B*27:05 bound peptides of the IR > 1.5 subgroup showed a statistically significant MW increase of 56 Daltons relative to the corresponding ligands strongly enriched in the presence of ERAP2 (Figure II.6, right panel).

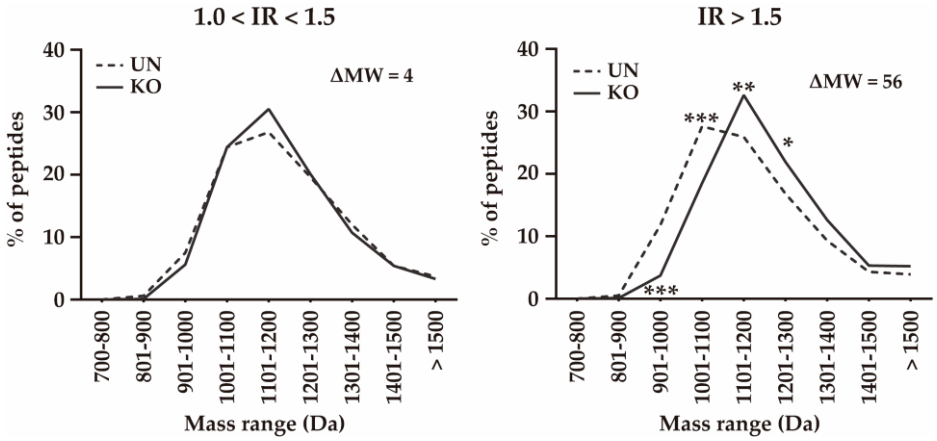


Figure II.6: Molecular weight (MW) of HLA-B*27:05 ligands from ERAP2$^{+/+}$ (UN) and ERAP2$^{-/-}$ (KO) cell clones. Peptides from UN or KO cells were subdivided based on their ion peak intensity rate (IR) according to the following conditions: 1.0 < IR ≤ 1.5 (left panel) and IR > 1.5 (right panel) and the equivalent subgroups were compared. The differences in the mean MW (ΔMW) of the peptides in each subgroup from both UN and KO cell clones are shown. Significant P values: *, P < 0.05; **, P < 0.001; and ***, P < 0.0001 are indicated in the corresponding MW range, respectively.

3. HLA-B*27:05 Ligandome in Absence of ERAP2 is Enriched in N-Terminal Basic Residues

The residue frequencies at different peptide positions of HLA-B*27:05 ligands identified in ERAP2$^{+/+}$ and ERAP2$^{-/-}$ cells were analyzed. No differences were found in the pairwise comparison of 1.0 < IR ≤ 1.5 subgroups from UN and KO cell clones at the N-terminal (P1) position. In contrast, comparison of IR > 1.5 subgroups from UN and KO cell clones reveals a statistically significant increase in the frequencies of Arg and Lys, and a significant decrease in the frequencies of Ile, Ser and Val in the absence of ERAP2 (**Figure II.7**).

4. Qualitative and Quantitative Enrichment of HLA-B*27:05 Ligands with Gln and Lys at Anchor Motif P2 Peptide Position in ERAP2$^{-/-}$ Cells

The residue frequencies at anchor motif P2 peptide position of HLA-B*27:05 ligands from ERAP2$^{+/+}$ and ERAP2$^{-/-}$ cells were analyzed. As previously described with the N-terminal (P1) position, no differences were found in the pairwise comparison of 1.0 < IR ≤ 1.5 subgroups from UN and KO cells; however, in the comparison of IR > 1.5 subgroups a statistically significant increase in the frequencies of Gln and Lys with concomitant decrease in the frequency of Arg at P2 peptide position were found in the absence of ERAP2 (**Figure II.7**, P2).

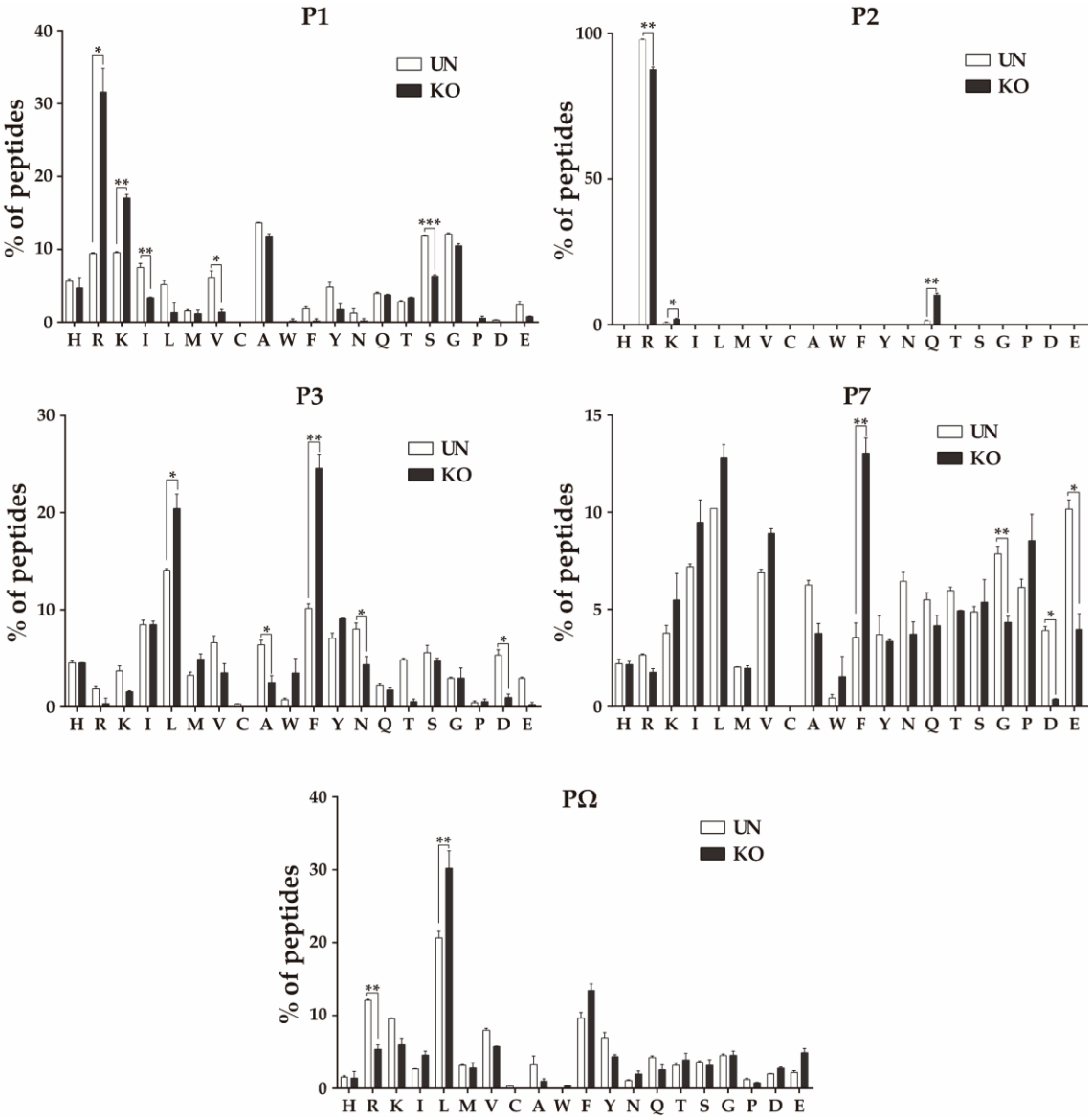


Figure II.7: Effect of ERAP2 depletion on the individual residue usage from HLA-B*27:05 ligands. Percentage of residue frequencies at the indicated positions in each individual panel among the peptides with IR > 1.5 from UN (open bars) and KO (closed bars) cell clones. These subgroups also include the peptides only found in UN or KO cell clones. Significant P values: *, P < 0.05; **, P < 0.001; and ***, P < 0.0001 between UN and KO cell clones are indicated, respectively.

Also, LC-MS signal intensities of shared ligands with Gln and Lys at anchor motif P2 peptide position were statistically stronger in the absence of ERAP2, relative to the same peptides identified in the presence of ERAP2 when compared with shared ligands with Arg at the P2 position (**Figure II.8**). No differences were found between shared ligands with Gln and Lys at the anchor motif position when peptide signal intensities from ERAP2$^{+/+}$ and ERAP2$^{-/-}$ cells were compared (**Figure II.8**).

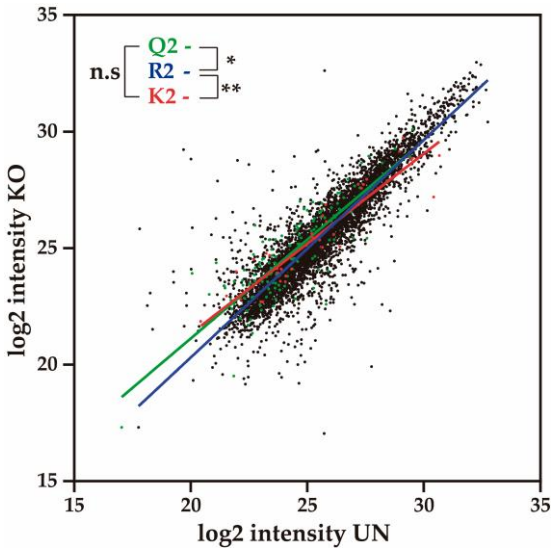


Figure II.8: Quantitation of the shared HLA-B*27:05 ligandome from ERAP2$^{+/+}$ (UN) and ERAP2$^{-/-}$ (KO) cell clones. Scatter plot of the mean LC-MS signal intensities of the HLA-B*27:05 ligandome from UN and KO cell clones. Peptides with Arg (R), Gly (Q), and Lys (K) at P2 position are colored in blue, green, and red, respectively. Linear regression of peptides with Arg, Gly and Lys at P2 position are colored in blue, green and red, respectively. Significant P values: *, $P < 0.05$; **, $P < 0.001$; and ***, $P < 0.0001$ between linear regression from UN and KO cell clones are indicated, respectively.

5. Qualitative Differences in Residue Frequencies at P3, P7 And PΩ, But Not at P1 Peptide Position of HLA-B*27:05 Ligands from ERAP2$^{+/+}$ and ERAP2$^{-/-}$ Cells

As previously described for the P1 and P2 positions, no differences were found in the pairwise comparison of $1.0 < IR \leq 1.5$ subgroups from ERAP2$^{+/+}$ and ERAP2$^{-/-}$ cells at any positions analyzed. When $IR > 1.5$ subgroups were compared, only the frequencies of some residues at P3, P7 and PΩ peptide positions were statistically significant. Leu and Phe were increased and, conversely, Ala, Asn and Asp were decreased in the absence of ERAP2 when P3 peptide position was assessed (**Figure II.7**, P3). At P7 peptide position, Phe was increased, whereas Asp, Gly and Glu were decreased in the ERAP2$^{-/-}$ cells (**Figure II.7**, P7). Analysis of the PΩ peptide position shows an increase in Leu and a decrease in Arg when ERAP2 is absent (**Figure II.7**, PΩ).

6. Influence of ERAP2 in the Hydrophobicity of HLA-B*27:05 Ligands

The hydrophobicity at the five relevant positions (P1, P2, P3, P7, and PΩ) of HLA-B*27:05 ligands modified by ERAP2 expression was measured with the GRAVY scale. A significant increase in the hydrophobicity at the N-terminal position peptide and a decrease in the hydrophobicity at P3, P7 and PΩ positions were found in ERAP2$^{-/-}$ cells (**Figure II.9**).

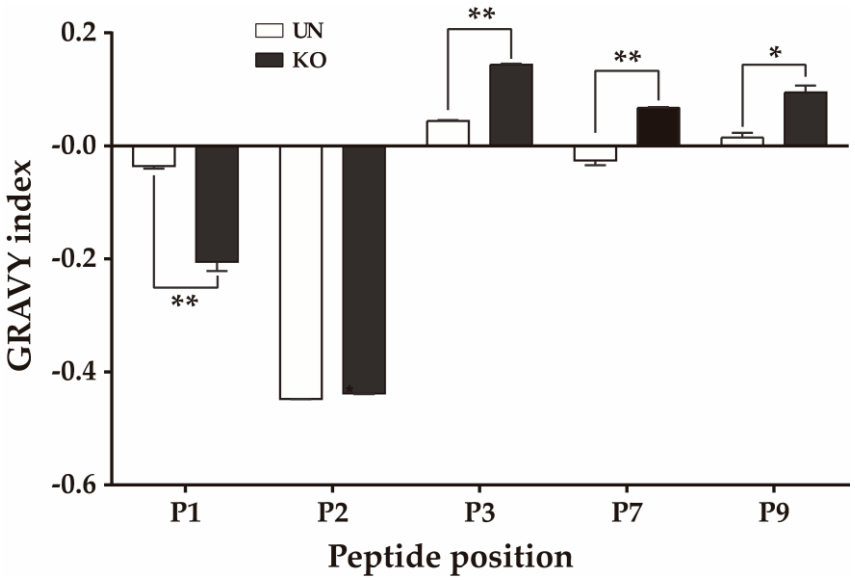


Figure II.9: Hydrophobicity of HLA-B*27:05 ligands from ERAP2$^{+/+}$ (UN) and ERAP2$^{-/-}$ (KO) cell clones. Comparison of hydrophobicity measured with the grand average of hydropathicity (GRAVY) scale (ProtParam tool, ExPASy Proteomics Server, available at www.expasy.ch) at the indicated position among the HLA-B*27:05 ligands with IR > 1.5 from UN (open bars) and KO (closed bars) cell clones. Significant P values: *, P < 0.05; **, P < 0.001; and ***, P < 0.0001 between UN and KO cell clones are indicated, respectively.

Discussion II

Since ERAP2 is a risk factor for the development of AS, in this study the role of ERAP2 aminopeptidase in shaping HLA-B*27:05 peptidome with CRISPR/Cas9 genome engineered ERAP2$^{-/-}$ cell lines was analyzed. The influence of ERAP2 in HLA-B*27 peptidome was already assessed in two previous studies using unrelated ERAP2$^{+/+}$ and ERAP2$^{-/-}$ lymphoblastoid cell lines (LCLs), showing a shift towards higher MW from ligands and an increment in peptides with N-terminal basic residues in the absence of ERAP2 **[181, 182]**. Here, the MW difference of HLA-B*27:05 ligands associated with ERAP2 was confirmed. Indeed, it was found a higher average mean (56 Daltons in CRISPR/Cas9-engineered ERAP2$^{-/-}$ cells versus a mean of 28 Daltons previously reported with LCLs **[181, 182]**). Additionally, enrichment of peptides with N-terminal basic residues in the absence of ERAP2 was detected in the mass spectrometry analyses using ERAP2-edited cell clones as previously described using ERAP2-positive and -negative LCLs **[181, 182]**. In addition, new roles for ERAP2 in the generation of HLA-B*27:05 ligandome were identified.

The expression of ERAP2 affects P1 residue usage by decreasing peptides with Ile, Ser and Val at this peptide position. These changes in P1 residue usage were associated with an increased hydrophobicity at the N-terminal position of HLA-B*27:05 ligands. For HLA-B*27:05 ligands, the P1 residue is an auxiliary anchor motif whose lateral chain interacts with different residues of the A pocket **[193]**. A previous study using poly-A analogs containing the Arg2 anchor motif and different residues at P1 position concluded that Arg was the amino acid with the highest contribution to HLA binding, Ile and Lys had a neutral contribution and Ser and Val were detrimental residues at P1 peptide position **[194]**. Thus, the results obtained with ERAP2$^{-/-}$ cell clones suggest that this aminopeptidase favors the generation of ligands with low affinity binding to HLA-B*27:05 class I molecules.

The interaction of P2 residue with the B pocket is the major contribution to B*27:05 binding, thus defining this position as an anchor motif (SYFPEITHI database: www.syfpeithi.de **[150]**). Arg is overwhelmingly the most utilized amino acid at P2 peptide position, with residual presence of Gln **[195-197]** and Lys **[198, 199]** at this anchor position.

Previous computational simulations have suggested a similar binding conformation between P2-Arg and P2-Lys peptides to the B pocket of the HLA-B*27:05 molecules, but slightly different from the P2-Gln peptides [198]. In the simulation, the hydrogen bond profiles were energetically and geometrically enhanced for Arg at P2 position, relative to Gln or Lys amino acids [198]. Also, binding analysis with the same peptide sequence and different changes at P2 position concluded that the HLA affinity was substantially lower for peptides with Lys or Gln as anchor motifs than the respective P2-Arg ligands [198]. Here we found that the absence of ERAP2 increases, in both quantity and quality, the frequencies of Gln and Lys amino acids at P2 anchor motif positions, thus indicating that HLA-B*27:05 in ERAP2$^{-/-}$ cells appear to be more tolerant to the loss of interaction in its B pocket.

In the absence of ERAP2, the frequencies at P3 peptide position of Ala, Asn and Asp decreased and, conversely, those of Leu and Phe increased, with a subsequent reduction in the hydrophobicity of HLA-B*27:05 ligands at this position. Previous studies using poly-A analogs containing the Arg anchor motif and different changes at P3 position concluded that Leu and Phe residues were favored in the interaction with the D pocket of HLA-B*27:05 class I molecules, thus increasing HLA binding. Moreover, Ala and Asn presented a neutral contribution, whereas Asp was detrimental at the P3 peptide position [194, 200]. Thus, in ERAP2$^{-/-}$ cell clones, the interactions with the D pocket seems to be more relevant for HLA-B*27:05 binding than in ERAP2$^{+/+}$ cells. The lateral chain of P7 residue interacts with the E pocket of HLA-B*27:05 class I molecules. Unfortunately, only one of the four amino acids for which the frequencies were different between ERAP2$^{+/+}$ and ERAP$^{-/-}$ cells identified in the screening was previously analyzed by using poly-A analogs [200], so no information about ERAP2 influence in this position could be obtained.

The lateral chain of PΩ residue interacts with the F pockets of HLA-B*27:05 class I molecules. At PΩ, the frequency of Leu is increased and, concomitantly, the frecuency of Arg decreased when *ERAP2* is absent. In two previous studies using poly-A analogs at PΩ residue, the relative binding affinity of Leu residue increased fourfold [194] or eightfold [200] that the one of Arg residue.

This scenario is compatible with the interactions at the F pocket, which are greater when ERAP2 is not available or absent. The overall picture of the results obtained is consistent with a remodeling of interactions between the different studies of ligands with different pockets of HLA-B*27:05 class I molecules influenced by ERAP2. The absence of this aminopeptidase increases the presence of ligands with a minority of canonical P2 residues, which presents low interactions in the B pocket but are compensated by the contribution of otherwise weak interactions between the auxiliary anchor motifs P1, P3 or PΩ to the A, D and F pockets, respectively.

The widely spread ERAP2 deficiency (with an incidence of ~25% of the human population) has a protective role for HLA-B*27 in ankylosing spondylitis **[174, 201]**. The qualitative and quantitative changes in the HLA-B*27:05 ligandome detected in ERAP2$^{-/-}$ cells may be relevant for studying the association between ERAP2 and this inflammatory type of arthritis in the context of the arthrogenic peptide hypothesis. The cellular peptides specifically generated by ERAP2, herein identified, may be tested as candidates for being arthrogenic peptides cross recognized by T cells. Another possibility will be that the quantitative differences in ERAP2 influence detected in multiple HLA-B*27:05 ligands may alter the equilibrium between immunogenicity and tolerance. Such a misbalance would trigger the inflammatory T cell response, expanding the universe of peptides to be tested as possible arthrogenic peptides.

Supplementary Figures II

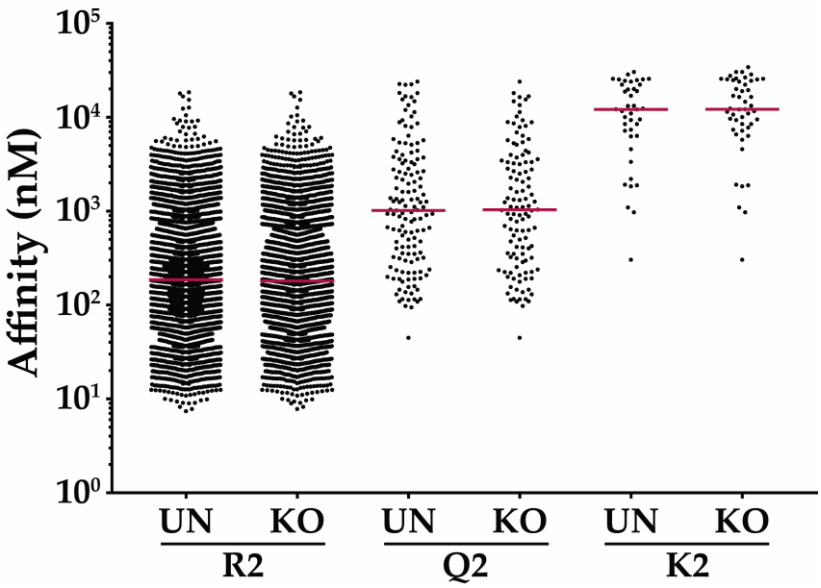


Figure SII.1: Theoretical affinity of HLA-B*27:05 ligands that were detected by MS/MS analysis. Theoretical affinity of identified peptides from UN and KO cell clones. Medians are indicated as continuous red lines.

Part B:

Laser Bioprinting Techniques as a New Approach for the Study of the Immune System

Chapter III: Fluorescence enhanced BA-LIFT: Development and Validation

Introduction III

1. Bioprinting Techniques

Bioprinting is a 3D fabrication technology that allows precise dispensing of cell-laden biomaterials and which can be applied to a variety of medical and biological fields. In bioprinting techniques, both cells and biomaterials are deposited with micrometer precision to form tissue like structures, requiring a special technical approach compatible with depositing living cells. Bioprinting allows accurate control of cell distribution and high-resolution cell deposition, being scalable and cost effective. For this reasons, work on the development and subsequent applications of bioprinting has greatly increased in recent years [202-205].

The three major bioprinting techniques used to date are inkjet, extrusion, and laser-assisted bioprinting, each having its specific strengths, weaknesses and limitations. Inkjet bioprinting was the first technology developed [207] and consist of a bioink (which consists of a polymer solution with cells) that is stored in a cartridge. The bioink is printed through a printer head that is deformed by a thermal or piezoelectric actuator and squeezed to generate droplets of a controllable size (Figure III.1 A). Inkjet bioprinting has three main advantages: **1)** low cost**; 2)** high printing speeds due to the use of multiple printing heads; and **3)** relatively high cell viability [208-210]**.**

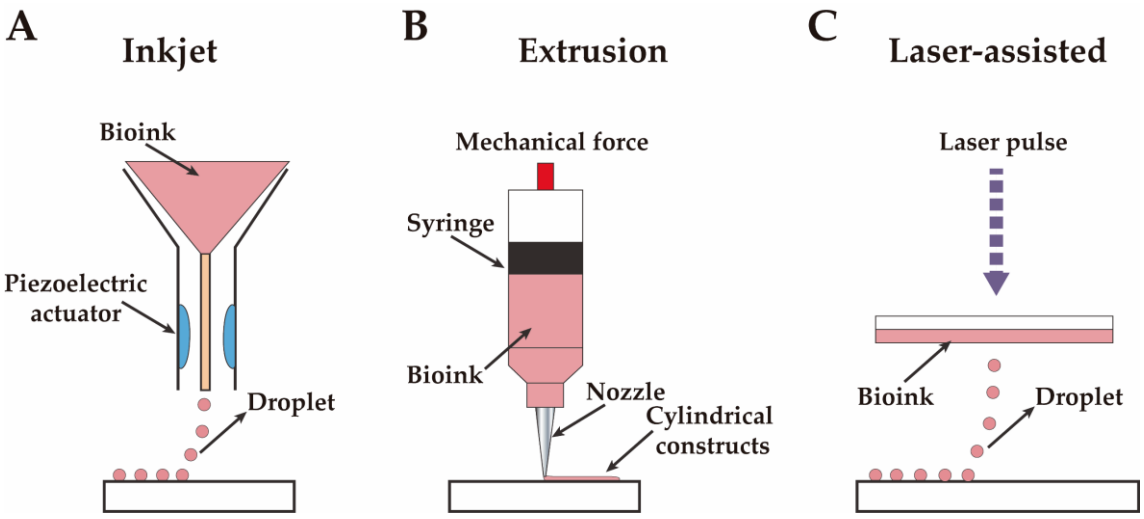


Figure III.1: Main bioprinting techniques. A) Inkjet bioprinting consists in the deposition of bioink droplets using a piezoelectric actuator**. B)** Extrusion uses mechanical forces to deposit a continuous cylindrical steam of bioink**. C)** Laser-assisted bioprinting uses an energy-absorbed donor layer that is vaporized after laser stimulation, thus ejects the hydrogel towards the acceptor layer underneath. Adapted from [206].

Nevertheless, this technique has major drawbacks: **1)** the printer heads cannot squeeze out highly viscous materials, and so are unable to work with bioinks with high cell density; **2)** easy clogging of the printer head [211-214]; **3)** a settling effect, that consists of cells settling in the cartridge, increasing the viscosity of the bioink and clogging the printer head [211, 212]; and **4)** low cell deposition resolution (> 100 μm accuracy).

Extrusion bioprinting is a modification of inkjet bioprinting which uses either an air force pump or a mechanical screw plunger to dispense the bioink (**Figure III.1 B**). This technique can print uninterrupted cylindrical lines, with almost all types of varying viscosity hydrogels and high cell-density aggregates. Nevertheless, extrusion bioprinters also expose the cells to a greater mechanical stress that is thought to reduce cell viability [215, 216]. The main disadvantages of extrusion printing are the elevated cost, the clogging of the nozzle and, more importantly, the high probability of sample contamination, which makes it necessary to sterilize the printing equipment.

Laser-assisted bioprinting was originated from laser-direct write and laser-induced transfer technologies. Compared to inkjet and extrusion bioprinting, laser-assisted can avoid the direct contact between the dispenser/cartridge and the bioinks, thus avoiding mechanical stress to the cells and resulting in high viability and lower contamination probabilities (**Figure III.1 C**). Laser-assisted bioprinting can print highly viscous bioinks, a wide variety of cell lines, and is faster and more reliable than the other bioprinting techniques.

2. Laser Induced Forward Transfer Bioprinting Technique

Laser Induced Forward Transfer (LIFT) is one of the most promising technique for biofabrication [215], suitable for printing a wide range of cell-laden bioinks [217-223].

Most of these applications use an approach based on a setup comprising a transparent donor substrate coated with an energy absorption layer and the material to be printed. Laser pulses are absorbed by the interface between the intermediate layer and the substrate, thus inducing the formation of a jet and transferring droplets of the biomaterial from the donor towards the acceptor substrate placed in parallel and separated at a certain gap distance [224].

For biological applications such as cell printing, the energy absorption layer can be made of: **a)** a thin layer of a metal (such as titanium, titanium oxide, silver or gold [218, 219, 225-228]); **b)** a hydrogel layer [229-231]**; c)** a polyimide layer [232-235]. With the latter approach, the laser pulse is absorbed by the substrate-polyimide interface, vaporizing a part of the polyimide layer and creating a vapor bubble that pushes and plastically and elastically deforms the remaining polyimide. The mechanically formed blister pushes the thin liquid layer situated on the polyimide, thus creating a jet and the corresponding droplet in a process called Blister-Actuated LIFT (BA-LIFT) [235]. The physics of both blister creation and jet droplet formation in BA-LIFT has been widely studied [233, 235-238].

Polyimide is a material widely used to take advantage of its excellent structural, electrical and thermal properties for applications that range from electronics to medicine [239]. Moreover, polyimide has high biocompatibility and has been used as a bio-inert material for implants, sensors and cell culture [240-244]. Like other polymers, polyimides have low wettability properties, so it requires surface treatment when high wettability is required [239]. Typical surface treatments used are prolonged UV exposure, plasma surface activation and chemical etching. One of the main benefits of using polyimide layers as an energy absorption layer is its transparency to visible light. This allows to combine conventional and fluorescence microscopy in the same optical path for the eventual cell detection and transfer.

Fluorescence microscopy has arisen as a powerful tool for cell and molecular biology, allowing direct visualization of the inner workings of physiological processes, as well as for identification, visualization and tracking of specific cell populations within a complex sample. The availability of large number of fluorescent probes and proteins, high performance detectors with high sensitivity and a wide variety of light sources and filters allows the integration of complete systems for faster and more precise fluorescence systems [245]. The combination of fluorescence imaging, together with the optical properties of the polyimide layers, enables to track cell populations within a sample both before and after the transfer process. Moreover, viability markers or dyes can be used to track cell viability *in situ* during the whole process.

Single cell analysis technologies have revealed heterogeneity between individual cells of the same cell type within a tissue. Detailed characterization of cell genome, transcriptome [246], proteome and metabolome [247] enables high-throughput profiling of cell characteristics that allows the definition of the functions that will occur both in normal and pathological processes. Although several methods have been developed for isolation of a single cell, there are limitations to the ability to obtain a single cell [248] from complex samples, from cells that do not express certain surface markers or to discern the efficiency of cell isolation. Therefore, BA-LIFT is emerging as a promising technique for isolating individual cells in a complex sample of previously enriched cell populations (by microfluidics or limiting dilution methods), in order to perform a deeper characterization.

This work aims to describe an original approach, called Fluorescence enhanced BA-LIFT (FE BA-LIFT), for single cell isolation using a BA-LIFT system with a commercial polyimide layer which is supplied in the form of an adhesive tape that can be easily fixed onto the donor and acceptor glass slides. The polyimide layer allows the implementation of a conventional and fluorescence microscopy module to enhance the potential of single cell detection and isolation properties of the system. Also, the printability maps for two biocompatible and widely used hydrogels (sodium alginate and methylcellulose) are presented, thus providing a general overview of the printing results related to the FE BA-LIFT parameters. With the help of printability maps, the optimum parameters in terms of the quality and size of the printed droplets can be obtained. Cell traceability, single cell isolation and viability has been evaluated using optimum transfer conditions through a combination of fluorescence markers and propidium iodide staining of the cells involved in the assessments. Moreover, the system's biological functionality was proven with two applications that required selection and detection using fluorescent markers. First, CRISPR/Cas9 genome edited natural killer (NK) cells were isolated, and their activation, individually and in groups, was tracked after printing in stimulation media. Second, the formation of colonies after transferring fluorescent marked hematopoietic stem and progenitor cells (HSPCs) from the bone marrow of a mouse was assessed after printing them onto methylcellulose hydrogels.

Materials and Methods III

1. Configuration of the FE BA-LIFT Bioprinting Device

BA-LIFT bioprinting device

The BA-LIFT bioprinting system comprises a laser source that emits pulses of 1.3 ns at 355 nm wavelength at a rate of 1,000Hz (Crylas, FTSS355-Q4). The pulse energy ranges from 1 μJ to 23 μJ, controlled by a half-wave plate and a beam polarizer. The laser beam is focused on the sample using a 10× microscope objective. A positioning X-Y axis system allows the donor and acceptor substrates to move independently of each other. The manual vertical movement (Z-axis) allows the samples to be placed on the focal plane of the vision system and the laser.

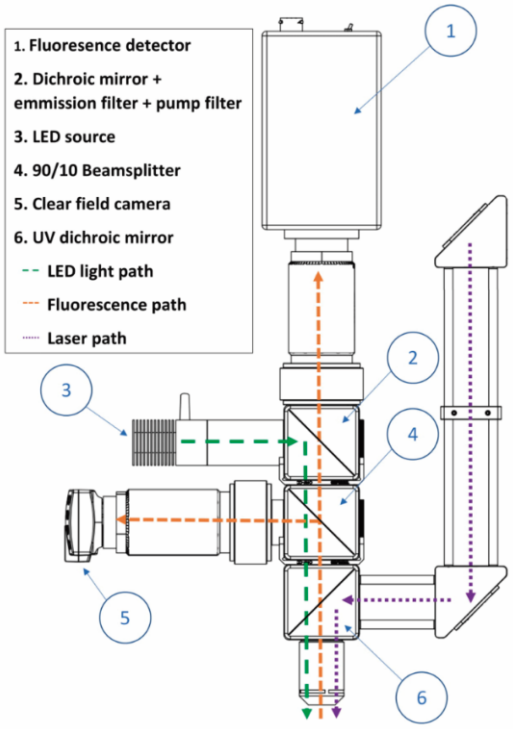


Figure III.2: Layout of the device used for FE BA-LIFT. All the elements included in the setup are shown, together with all optical paths followed by the light, fluorescence setup and laser.

Conventional microscopy and fluorescence vision system

Both conventional microscope imaging and fluorescence vision systems were implemented on the bioprinter, allowing to monitor the cells simultaneously with the transfer process. **Figure III.2** depicts the vision system, which consists of an LED source from which light is filtered through an excitation filter (**Figure III.2 '6'**) to select the wavelength required for each fluorophore. The filtered light goes through a set of dichroic mirrors, allowing the transmission of visible light and acting as a mirror for UV light. With this setup, the laser beam can be focused through the same microscope objective used by the imaging system.

The signal obtained from the sample goes first through a beam splitter (**Figure III.2 '4'**), dividing the light into two paths: 10% of the light goes to the clear field image camera (**Figure III.2 '5'**), and the remaining 90% passes through the emission filter (**Figure III.2 '2'**) before reaching the fluorescence detector (**Figure III.2 '1'**).

Donor and acceptor slides

BA-LIFT is performed by printing material from a transparent substrate called donor to another substrate called acceptor (**Figure III.3**). The donor substrate was made with a 25mm x 25mm microscope slide; whereas the acceptor substrate was changed depending on the performed experiment: **a)** viability and single cell isolation experiments were carried out using a 25 mm x 25 mm soda lime microscope slide as an acceptor substrate; **b)** proliferation and activation assays were carried out using a 96 well plate filled with culture medium as an acceptor substrate**,** and **c)** CFU assays were carried out using a 35 mm culture dish filled with 1.27% methylcellulose as an acceptor substrate.

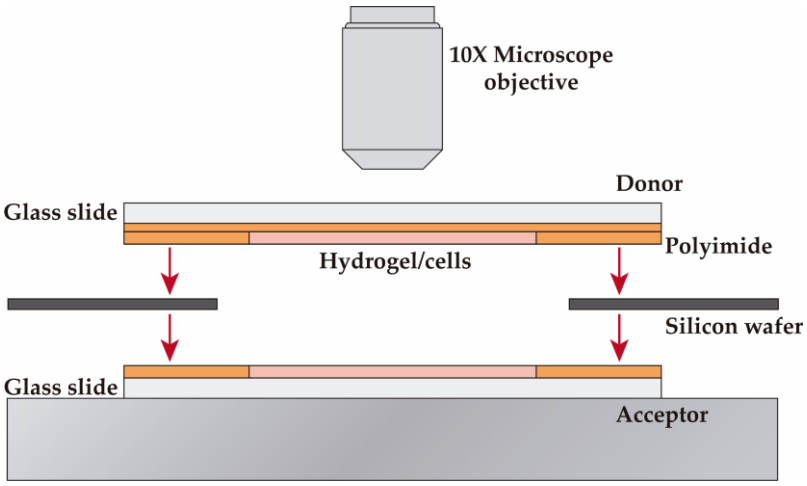


Figure III.3: 2D section representation of the donor and acceptor slides used in FE BA-LIFT, showing the wells made of polyimide. The acceptor substrate was changed depending on the experiment performed, as explained above.

On the donor, a layer of thin polyimide tape (Caplinq, PIT0.5S-UT/19) was applied, which functioned as the energy absorption layer; this is where the blister is formed. Afterward, a second thicker layer (~70 μm) of polyimide tape (Caplinq, PIT1S/9) was applied on top of the previous layer; subsequently, a 14mm diameter circle was cut out and removed from this layer, creating a well that contain the hydrogel.

Plasma surface activation and cell-laden hydrogel spreading

To improve the wettability properties of the donor substrate polyimide layer, the surface was activated with a plasma treatment. The plasma treatment consisted in exposing the surface to a plasma electrically generated by microwave irradiation. Samples were placed inside a small glass desiccator chamber connected to a vacuum pump and taken to a pressure of ~133 Pa. The desiccator was exposed to microwave radiation until plasma was formed and the surface ionized. 15 µl of the cell-laden hydrogel (8×10^6 cells/ml) was placed on the donor using a blade coater; then, the donor was turned over and placed on the FE BA-LIFT device for cell printing.

2. Hydrogels Used to Create the Printability Map and for The Colony Forming Units (CFU) Assays

Printability maps were created by printing two hydrogels: sodium alginate and methylcellulose. Sodium alginate (Sigma-Aldrich, W201502) was dissolved in a saline solution (NaCl 0.9%) to create a 4% stock; then the stock was diluted with RPMI 1640 culture medium (Lonza, BE12-702F) to obtain the final concentrations of 1.5% and 2%. A methylcellulose 3% stock solution (R&D Systems, HSC001) was mixed with IMDM culture medium (Gibco, 1244-053) to obtain the final concentrations of 1.5%, 1.2%, 1%, 0.8%, 0.5% and 0.3%. For CFU assays, a methylcellulose 3% stock solution (R&D Systems, HSC008) was mixed with IMDM culture media plus 2% FBS to a final concentration of 1.27%. The hydrogels were stored at 4°C, and used at room temperature.

3. Design and Generation of IFN-γ Fluorescence Reporters

To perform the FE BA-LIFT validation through activation assays, a human cell with a consistent activation marker had to be used. For that purpose, the human IL-2 independent NK cell line NK-92® MI (ATCC®, CRL-2408™) and the activation marker IFN-γ were chosen to track cell activation after the printing process. Insertion of the *eGFP* coding sequence at the end of *IFN-γ* exon 4 was carried out using CRISPR/Cas9 genome editing with two different approaches: **1)** by generating an IFNγ-eGFP fusion protein, and **2)** by bicistronic expression of the eGFP together with IFN-γ.

Characterization of human *IFN-γ* genomic *locus* and CRISPR/Cas9 strategy design

Sequence and genomic *locus* distribution of human *IFN-γ* (Chr12:68,154,768 - 68,159,740, GRCh38/hg38 database) was obtained from the UCSC Genome Browser [87] (ENCODE consortium [88]), and 5′ and 3′ UTRs, coding exons and introns were identified and depicted using Snapgene® (from GSL Biotech; available at snapgene.com).

Exon 4 and 3'UTR were analyzed using the CRISPR/Cas9 design tool from Benchling (Biology software, 2019. Retrieved from benchling.com) and the best suitable crRNA (hIFNγ_KI_stop; sequence: GTCGAAGAGCATCCCAGTAA; PAM: TGG; target: hIFNγ STOP codon; genomic location in GRCh38/hg38: Chr12: 68,155,353 - 68,155,372) was chosen following the parameters described in **M&M I.5** and in **Figure III.10**.

Transfection of NK-92® MI cells was carried out as described in **M&M I.5**, with the addition of the HDR donors.

IFNγ-eGFP fusion protein

In frame insertion of the *eGFP* coding sequence after *IFN-γ* exon 4 was performed by transfecting 1 µg of an HDR plasmid donor together with the CRISPR/Cas9 RNP. For HDR plasmid donor generation, a 2,000 bp gBlock® gene fragment (Integrated DNA Technologies, **Figure SIII.3**) comprised of a 615 bp 5′ homology arm, the *eGFP* coding sequence, and a 668 bp 3′ homology arm was synthesized. The leucine 158 codon CTG was changed for CTC and the phenylalanine 159 codon TTT was changed for TTC to avoid the binding of the RNP complex to the HDR donor. To allow the expression of IFNγ-eGFP fusion protein, the STOP codon (TAA) was removed from the HDR donor. The HDR donor was cloned into the Zero Blunt™ PCR vector (Thermo Fisher Technologies, K270020), following the manufacturer's protocol, thus generating the HDR plasmid donor.

Bicistronic expression of the eGFP together with IFN-γ

For bicistronic expression of the eGFP and IFN-γ, 500ng of a 903 bp ssODN donor was transfected together with the CRISPR/Cas9 RNP. The ssODN donor (Megamer™ ssDNA fragment, Integrated DNA Technologies, **Figure SIII.4**) consisted of a 60 bp 5′ homology arm, the 2A self-cleaving peptide with a glycine-serine-glycine spacer (P2A-GSG)

sequence, the *eGFP* coding sequence and a 60 bp 3′ homology arm. The same codon mutation as in the HDR donor was performed, but the *IFN-γ* STOP codon was not removed as the two proteins would be translated independently.

NK-92® MI-edited cells were analyzed by PCR 48 hours after transfection for the *eGFP* sequence insertion using the primers depicted in **Table III.1.** eGFP⁺ cells were sorted using FACS (**M&M III.4**) and then cloned. The clones were analyzed again by PCR to select those cells with homozygous insertion of the *eGFP* coding sequence. The selected cell clones were used for further analysis of IFNγ-eGFP expression and activation assays.

Table III.1: PCR primers used for *IFN-γ* genotyping and the expected sizes of the amplicons.

Primer	Sequence 5′ → 3′	Amplicon length		Genomic location, GRCh38/hg38
		Unedited	IFNγ-eGFP	
IFNg_Fw	TGTCCAACGCAAAGCAATACA	419 bp	Mk. I: 1133 bp	Chr12: 68,155,453 – 68,155,473
IFNg_Rv	CACAACCCATGGGATCTTGC		Mk. V: 1202 bp	Chr12: 68,155,055 – 68,155,074

4. Flow Cytometry Analysis and Cell Sorting

Intracellular IFN-γ staining of CRISPR/Cas9-edited NK-92® MI cells was performed 6 hours after stimulation with 50 ng/ml PMA plus 1 μg/ml ionomycin in the presence of 5 μg/ml brefeldin A (BFA). Stimulated cells were washed twice with 1X PBS and fixed with 4% paraformaldehyde (Electron Microscopy Sciences) for 12 minutes at room temperature in the dark, and then permeated with permeation buffer (1X PBS plus 1% saponin (Sigma- Aldrich, S7900) and 3% FBS) for 20 minutes at 4°C. The permeated cells were labeled with 2 μl of an α-hIFNγ-PE/Cy7 monoclonal antibody (Biolegend, 506518, clone B27). The edited NK-92 MI reporter cells were PMA/Io stimulated for 6 hours and GFP⁺ cells were sorted by FACS to further cloning.

Flow cytometry samples were acquired with a FACSCanto (Becton Dickinson, Franklin Lakes, NJ, USA) flow cytometer; whereas FACS samples were acquire with a FACSAria flow cytometer. The data was analyzed and represented using FlowJo™ software.

5. Confocal Microscopy

Confocal microscopy was used to assess the eGFP production of human IFN-γ reporter NK-92® MI cells. NK-92 *Mk. I* and NK-92 *Mk. V* were stimulated for 6 hours with 200 µl of α-MEM culture media supplemented with 50 ng PMA and 1 µg ionomycin. Cells were incubated 10 minutes with 1 µg/ml Hoescht 33342 (ThermoFisher Scientific, 62249), washed with 1X PBS and placed on a microscope slide for fluorescence analysis.

Confocal fluorescent images were obtained with a Leica TCS SP5 AOBS confocal scan head mounted on a DMI6000 inverted based microscope with an HCX PL APO CS 40×/1.25 and HCX PL APO 63×/1.4 objectives. Excitation at 474 nm was provided by an argon gas laser. The 525 nm green fluorescence was collected with a B filter (blue excitation BP 450-490, emission LP 515) and additional Nomarsky-DIC images were taken with LAS-AF software (version 2.6.3, Leica, Microsystems). Images were analyzed with Fiji ImageJ software [249].

6. Assessment of Cell Viability, Activation Capabilities and Proliferation After Printing Process

Single cell isolation using FE BA-LIFT

For single cell isolation experiments, a B lymphocyte that constitutively expresses eGFP (C1R-N1-85) and a T lymphocyte Jurkat cell line were used. 10^6 Jurkat cells were stained with 0.05 µg of an α-CD45/PE monoclonal antibody (Biolegend, 368510) for 30 minutes at 4°C in the dark. 8-10×10^6 C1R-N1-85 and Jurkat stained cells per milliliter were mixed with the hydrogel at a 1:1 ratio and spread over the donor layer. Cells were individually printed onto the acceptor substrate by the combination of the vision system and the laser setup.

Cell viability and proliferation assays

For viability assays, hydrogels with 6 µM propidium iodide (PI) were prepared both for the donor and the acceptor substrates. Non fluorescent cells in the donor substrate were printed, and fluorescent cells in the acceptor substrate were labeled as dead. For proliferation assays, C1R B cells and Jurkat T cells were printed directly into a 96 well culture plate containing 200 µl of RPMI 1640 culture media at fluences of 11J/cm^2,

5.7 J/cm², 3.1 J/cm² and 2.1 J/cm². Proliferation was assessed by means of optical microscopy tracking the cells every day for 4 days.

Cell activation assays

Cell activation assays were carried out using the human NK cell line with bicistronic expression of IFNγ and eGFP (*Mk. V*, **M&M III.3**). Cells were printed directly into wells of a 96 well culture plate containing 200 µl of stimulation media (α-MEM, Gibco 12561056, plus 50ng PMA and 1 µg ionomycin). Cells were incubated for 6 hours after printing and IFN-γ-eGFP production was assessed by fluorescence microscopy.

HSPCs isolation and CFU assays

Mouse bone marrow cells were obtained from C57BL/6 mice and stained with a biotin conjugated antibody cocktail consisting on an α-CD3ε (eBioscience, 13-0031-82, clone 145-2C11), α-CD4 (Southern Biotechnologies, 1540-08, clone GK 1.5), α-CD8 (eBioscience, 13-0081-85, clone 53-6.7), α-CD19 (eBioscience, 13-0193-85, clone eBio 1D3), α-CD220 (eBioscience, 13-0452-82, clone RA3-6B2), α-NKp46 (Biolegend, 137616, clone 29A 1.4), α-CD11b (BD Pharmingen, 553309, clone H1/70) and α-CD11c (BD Pharmingen, 553800, clone HL3) for 30 minutes at 4°C in the dark. Cells were washed twice with 1X PBS and then incubated for another 30 minutes at 4°C in the dark with streptavidin-PE to label lineage$^+$ (Lin$^+$) as red cells and lineage$^-$ (Lin$^-$) as non-fluorescent cells.

To isolate HSPCs, Lin$^-$ cells were printed onto a culture dish with 1.27% methylcellulose (R&D Systems, HSC008) and placed in the incubator at 37°C and 5% CO_2. Colonies, which reflected the number of CFUs, were counted at different days after the FE BA-LIFT process.

Results III

1. Bioprinting of Immune System Cells Using FE BA-LIFT

Printability map of sodium alginate and methylcellulose

The printing properties of the hydrogels were assessed at different fluence ranges to create a printability map. Characteristics analyzed were droplet shape, diameter and the number of satellite droplets produced during the printing. Printability maps were created for sodium alginate and methylcellulose. Data was taken by performing a total of 24 linear passes/sample, each pass being separated by 400 μm and made at a speed of 400 mm/s, therefore separating the droplets the same distance on the X- and Y-axis. A total of 18 images per laser fluence and hydrogel concentration were taken, and 6 to 8 droplets per image were measured and analyzed. Fluence varied from the maximum down to the limit at which no droplets were printed.

Tables **III.2** and **III.3** show the printability map of sodium alginate and methylcellulose, categorizing the quality of the droplets by a 4 color code: red indicates a bad shape with numerous satellite droplets (**Figure III.4 Ai**); orange indicates a bad shape with fewer satellite droplets (**Figure III.4 Aii**); light green indicates a good shape with some satellite droplets (**Figure III.4 Aiii**); and dark green indicates a good shape with no satellite droplets (**Figure III.4 Bi**). Measurement of droplet diameter was made for each set of experiments using an average of all data and defining the standard deviation as an indicator of repeatability. Laser fluence used varied from 6.5±0.5 J/cm^2 to 2.3±0.5 J/cm^2. **Figure III.4 A** and **SIII.1 s**how the results obtained using sodium alginate, taking into account that concentrations higher than 2% do not allow cell growth **[250]**, whereas, at concentrations below 1.5%, good results were not achieved. The parametric window when working with this hydrogel is quite narrow, as diameter measurements showed a big dispersion and satellite droplets appeared in most of the fluence conditions assayed. Methylcellulose can be used for cell culture at 1.27% concentration, as recommended by the manufacturer. In comparison with sodium alginate at its maximum accepted concentration, methylcellulose has a higher viscosity, so the transfer required higher fluences. Transfer of 1.5% methylcellulose droplets was only achieved at the highest fluence.

Table III.2: Sodium alginate printability map. Red indicates a bad shape, with numerous satellite droplets; orange indicates a bad shape with fewer satellite droplets; light green indicates a good shape with some satellite droplets; and dark green indicates a good shape with no satellite droplets.

Fluence (J/cm²)	Sodium alginate concentration	
	2%	1.5%
6.5		
6.3		
6.2		
6.1		
5.9		
5.8		
5.4		
5.1		
4.8		
4.2		
4.0		
3.7		
3.1		
2.5		
2.3		

Table III.3: Methylcellulose printability map. Red indicates a bad shape, with numerous satellite droplets; orange indicates a bad shape with fewer satellite droplets; light green indicates a good shape with some satellite droplets; and dark green indicates a good shape with no satellite droplets. Grey indicates the range of fluences, and hydrogel concentrations where no printing was achieved.

Fluence	Methylcellulose concentration					
J/cm²	1.5	1.2	1.0	0.7	0.5	0.3
6.5						
6.3						
6.2						
6.1						
5.9						
5.8						
5.4						
5.1						
4.8						
4.2						
4.0						
3.7						
3.1						
2.5						
2.3						

Nevertheless, it is easier to find a tendency in data for methylcellulose; at the same printing fluence, droplets are smaller as hydrogel concentration increases (**Figure III.4 B** and **SIII.2**). The same tendency is observed when the hydrogel concentration is fixed and the laser fluence is changed; droplets tend to decrease in size as the fluence reduces. Maximum repeatability was obtained at low fluences where smaller droplets were printed. The best results regarding shape, size and repeatability were obtained for methylcellulose, and so this hydrogel is proposed to be used for the study of cell viability after printing.

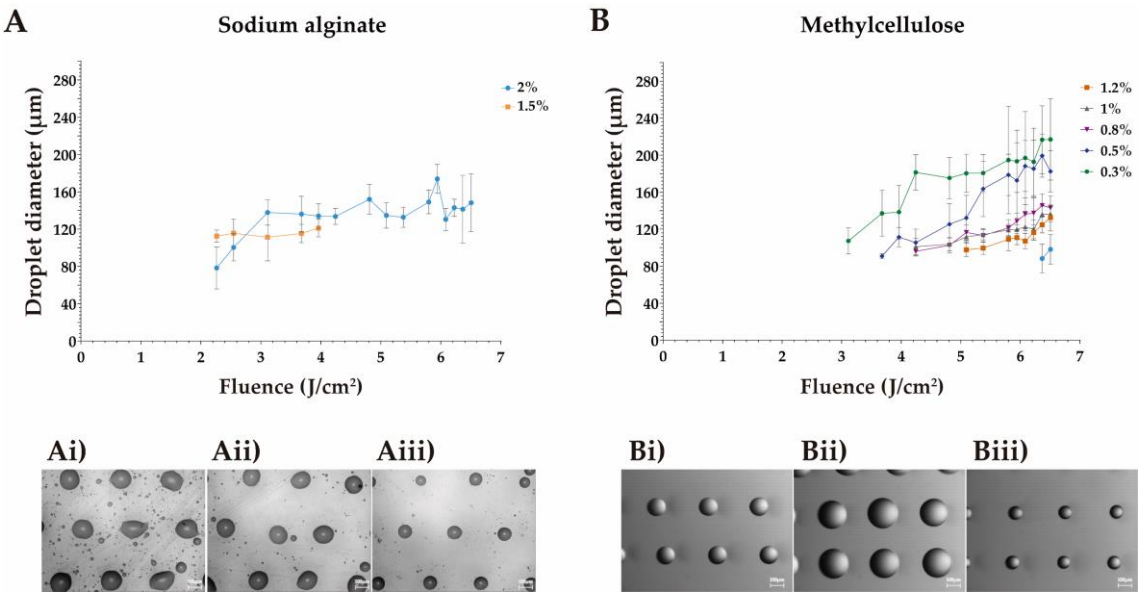


Figure III.4: Printability of sodium alginate and methylcellulose. A) Sodium alginate droplet diameters for all experiments. Examples of 2% sodium alginates at: **Ai)** 6.1J/cm², **Aii)** 5.1 J/cm² and **Aiii)** 4.8 J/cm². **B)** Methylcellulose droplet diameters for all experiments. Examples of methylcellulose droplets at: **Bi)** 1% at 6.1 J/cm², **Bii)** 0.5% at 5.4 J/cm², and **Biii)** 1.5% at 6.4 J/cm². Scale bars indicates 100 µm.

Viability and proliferation of printed cells

For viability assays, methylcellulose hydrogel (both on the donor and the acceptor substrate) was mixed with propidium iodide (PI), a fluorescent intercalating agent that binds to DNA between the bases with little or no sequence preference. Cell viability was assessed before (**Figure III.5 A**) and after the printing process (**Figure III.5 B, C** and **D**) through the vision system using conventional and fluorescence microscopy, thus allowing the identification of non-viable or dead cells during the FE BA-LIFT printing.

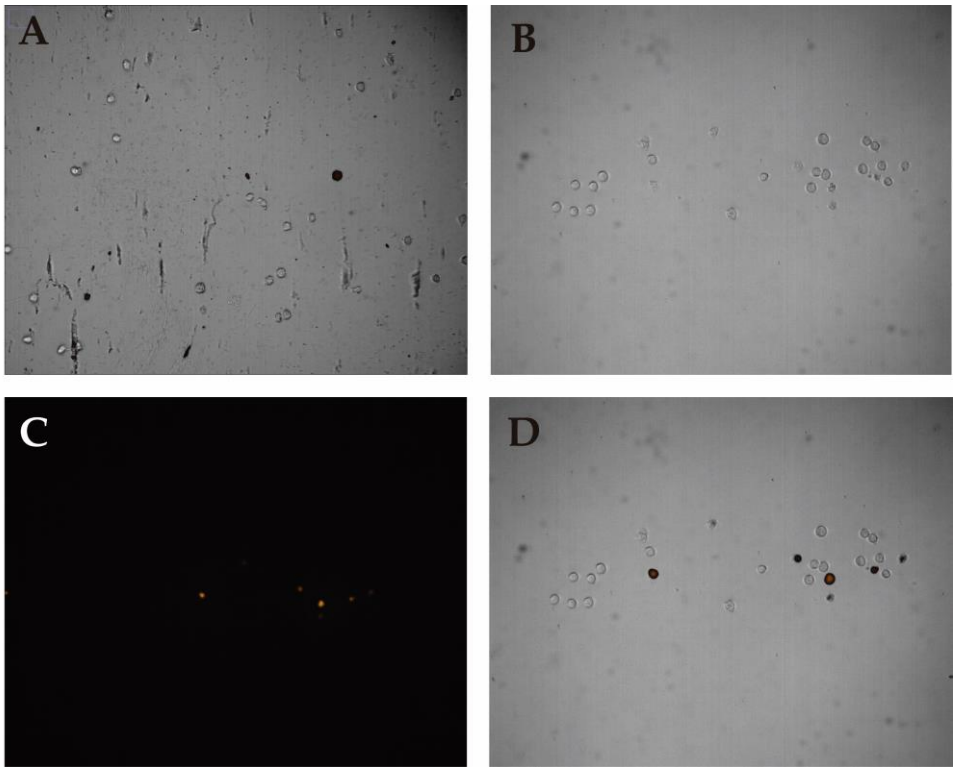


Figure III.5: Cell viability measurement using fluorescence enhanced (FE) BA-LIFT. A) Superposed microscopy image (conventional and fluorescence) of viable and dead cells in the donor substrate**. B)** Conventional microscopy image of transferred cells**. C)** Fluorescence image of transferred cells**. D)** Superposed image of transferred cells. All images were taken with the FE BA-LIFT system**.**

Viability assays were carried out by spreading 8-10×10^6 Jurkat T cells or C1R B cells on a 0.3% methylcellulose hydrogel plus 0.3 μM PI. Cells were printed at different fluences and, taking advantage of the vision system, all printed cells were considered in the counting process and viability calculation. **Figure III.6** shows the viability of the cells after the printing process for fluences of 6.1 J/cm^2, 5.7 J/cm^2, 5.4 J/cm^2, 4.8 J/cm^2, 3.4 J/cm^2, 3.0 J/cm^2, 2.5 J/cm^2 and 2.0 J/cm^2 at two laser focusing positions (P0 and P-480). At P0, no correlation was seen between viability and fluence; viability variation between experiments was up to 6%, making the printing process relatively repeatable but not ideal. Maximum and minimum viabilities were 84% and 68%, respectively. Printing results at the P-480 laser focusing position were very similar to those obtained at P0; nevertheless, there was a clear influence on viability. At all fluences tested viabilities were higher at P-480, with better repeatability; additionally, at fluences higher than 2.4 J/cm^2 viability seems not to be further affected by an increase in fluence. More repeatable results were obtained at the lowest fluences with a viability of 98%; the minimum viability obtained being 85%.

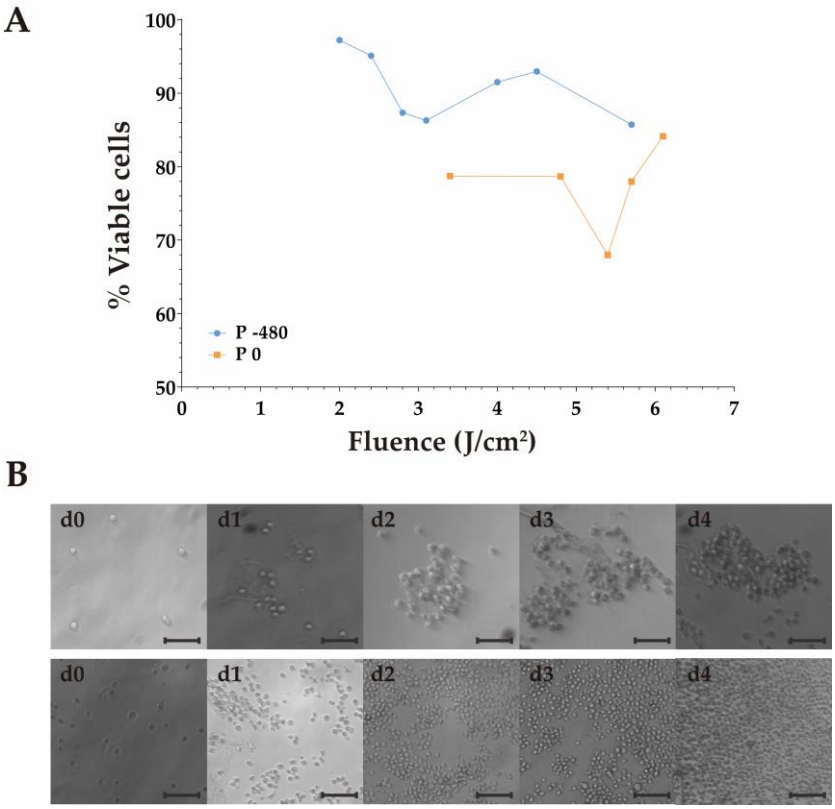


Figure III.6: Viability at P -480 and P0 focus positions and proliferation assays for Jurkat T cells and C1R B cells. A) Viability for different fluences at both working positions (P-480 and P0)**. B)** Measurement of cell proliferation. The upper panel depict C1R cells printed at 2.1 J/cm²; the bottom panel depict Jurkat cells printed at 5.7J/cm²**.** Scale bars indicate 100 μm.

It has been described that viability of printed cells depends on a combination of different parameters [225]. In the BA-LIFT approach cell viability does not seem to be greatly affected either by the donor hydrogel preparation or by the laser energy, although more experiments are needed to understand the residual cell death inherent in the transfer process at both laser focusing positions.

Cell proliferation studies were performed for both Jurkat T cells and C1R B cells at the laser focusing position where more reliable results were obtained. Cells were printed into a 96well plate filled with 200 μl of RPMI culture medium. Proliferation assays were performed at fluences of 11 J/cm², 5.7 J/cm², 3.1 J/cm² and 2.1 J/cm². **Figure III.6 B** upper panel shows an example of C1R cells printed at 2.1 J/cm² after 4 days of cell culture, and the bottom panel shows the proliferation of Jurkat cells printed at 5.7 J/cm² after 4 days of cell culture. Proliferation was observed for all fluences up to the third day; most of the cells printed at 11 J/cm² were dead at the fourth day, whereas at all other of fluences, cells continued proliferating.

The number of cells after culture only depends on the initial number of living cells printed into the well, so proliferation does not seem to be affected by the laser fluence.

Single cell isolation

Implementation of a vision system consisting of a bright field and fluorescence modules in the BA-LIFT printing device allows the identification and printing of cells that express specific surface markers or fluorescent reporters. Specific cell selection, combined with the capability to recognize viable cells before and after the printing process, enables the design of tools for developing an alternative sorting method to isolate specific cells within a complex population. Furthermore, the FE BA-LIFT device allows assessment of the survival rate of cells before and after the printing process, with a non-toxic staining method.

To assess BA-LIFT's capabilities to isolate single cells from a mixture of different cell populations, two cell lines were used: **1)** A C1R-N1-85 which constitutively expresses eGFP; and **2)** Jurkat stained with a monoclonal antibody against the surface receptor CD45 labeled with the PE fluorophore (α-CD45/PE) (**Figure III.7**). C1R-eGFP cells were first mixed with unlabeled Jurkat cells in a 1:1 ratio (**Figure III.7 A**); then, C1R-eGFP cells were mixed with the α-CD45/PE labeled Jurkat cells (**Figure III.7 B**). Cells can be identified within the focal depth of the microscope objective because, after placing the donor-acceptor setup, cells tend to deposit in a single focal plane relative to the lower layer of the hydrogel. In both cases, cells are easily identified and targeted, providing a suitable tool for the identification of specific cell lines within a complex sample.

Furthermore, the capability of the FE BA-LIFT system to isolate single cells within a population of mixed cells was tested. For effective isolation of single cells, a minimum separation of 50 μm between cells was necessary, as shorter distances result in printing of adjacent cells in the same droplet. The cell concentrations used (~8×10^6 cells/ml hydrogel) were enough to keep all cells at least 50 μm apart on the FE BA-LIFT donor surface (~14 mm diameter).

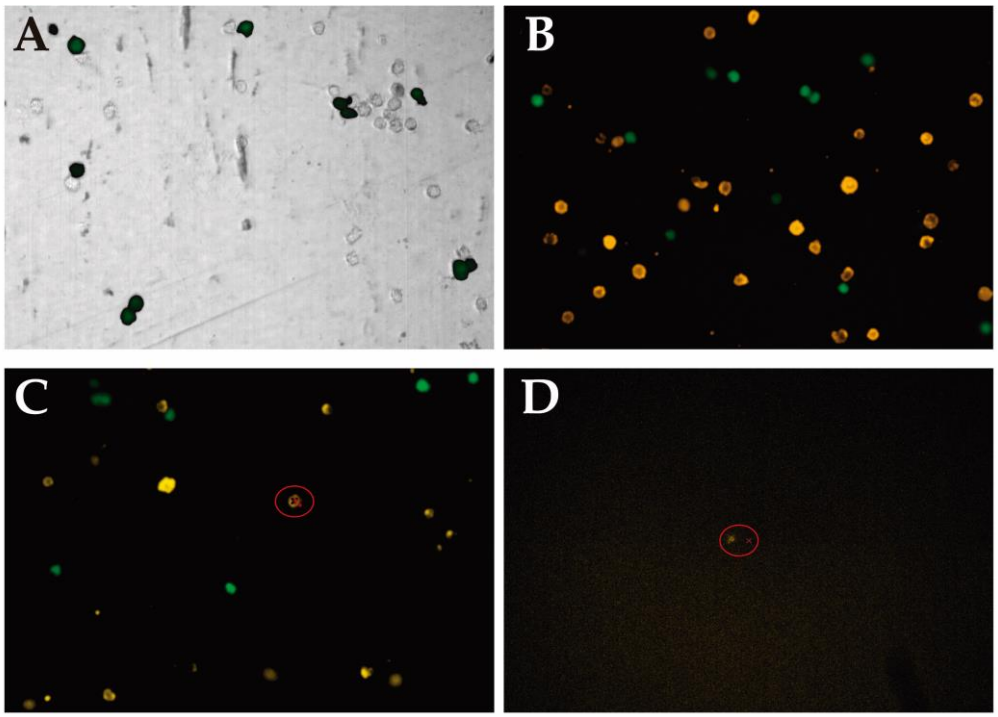


Figure III.7: Single cell detection and isolation using the FE BA-LIFT. A) Overlay of both conventional and fluorescence microscopy images of C1R-eGFP cells mixed with unlabeled Jurkat cells**. B)** Fluorescence image taken with the FE BA-LIFT system of C1R-eGFP cells mixed with α-CD45/PE labeled Jurkat T cells**. C)** $CD45^+$ cell selected to be printed with the FE BA-LIFT system**. D)** Single printed cell tracked with the same system used to select the cell on the donor substrate**.**

Single cells can be isolated along the whole donor surface, selecting the $CD45^+$ Jurkat T cells and printing them to the acceptor substrate (**Figure III.7 D**). The cell survived the printing process and retained its membrane integrity, allowing its tracking along the acceptor substrate.

NK cell printing and activation assays

The functionality of the cells after the printing process was assessed by a cell activation assay using a genome-engineered NK cell line. An NK-92 *Mk.V* cell line with a bicistronic expression of eGFP and IFN-γ (**M&M III.3**) was printed directly into a 96 well plate filled with 200 µl of stimulation culture medium (α-MEM plus 50 ng PMA and 1 µg ionomycin), and cell activation was measured by means of the eGFP fluorescence, which is similar to the IFN-γ production (**Figure III.8**). eGFP fluorescence was measured after 4 (**Figure III.8 D**) and 6 hours (**Figure III.8 F**), being most intense after 6 hours. **Figure III.8 A** and **B** shows a single NK-92 *Mk. V* printed cell both with conventional and fluorescence imaging, respectively.

PMA/Io activation favors cell aggregation, a phenomenon which was observed when a group of cells was printed, therefore indicating that the laser printing process does not affect NK cells' metabolic activity and activation capabilities.

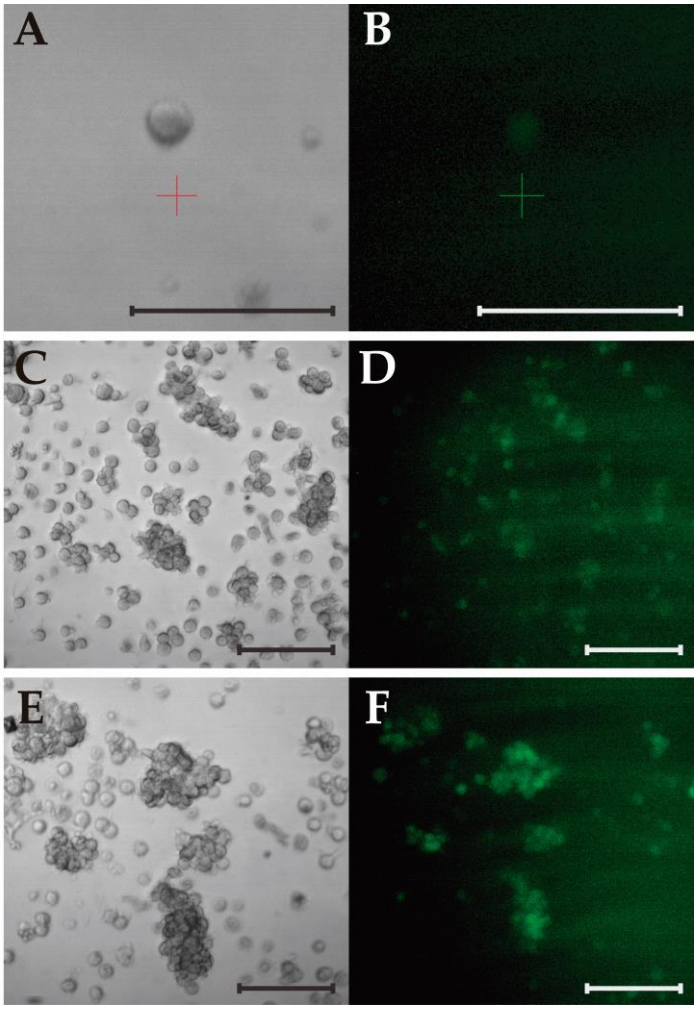


Figure III.8: NK cell activation tracked using the FE BA-LIFT system.
Conventional **(A)** and fluorescence **(B)** images of a single NK cell printed in a 96 well plate 4 hours after incubation with stimulation media. Conventional **(C)** and fluorescence **(D)** images of a group of NK cells 4 hours after incubation. Conventional **(E)** and fluorescence **(F)** images of a group of NK cells 6 hours after incubation. Scale bars indicate 100 μm.

HSPC survival and proliferation, and CFUs assays

The influence of the printing process on the differentiation potential of primary cells was assessed by printing HSPCs onto a 35mm diameter petri dish filled with 1.1 ml of 1.27% methylcellulose. Bone marrow cells from C57BL/6 were extracted and identified with molecular markers of differentiated hematopoietic cells as a Lin^- and Lin^+ cells (**Figure III.9**, **M&M III.6**).

Figure III.9 C depicts the merged picture of the bright (Figure III.9 A) and fluorescent (Figure III.9 B) images, where red cells correspond to Lin^+ cells and non-stained cells correspond to Lin^- cells. The cell marked with a blue circle was the HSPC selected for printing.

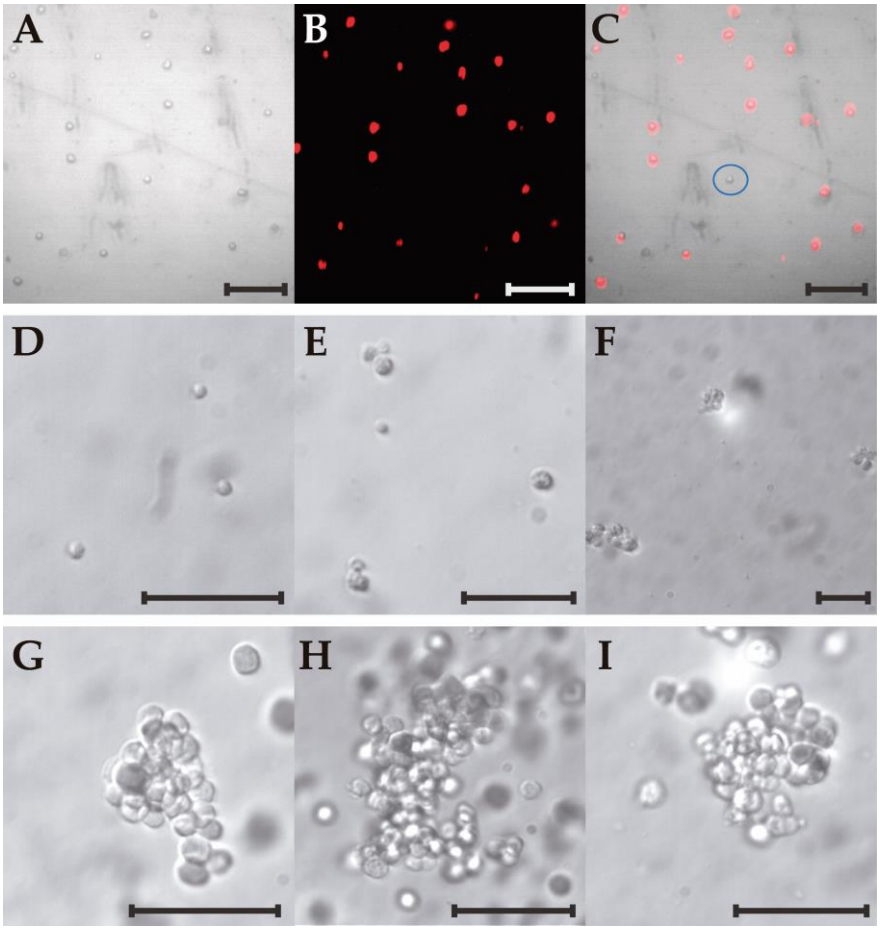


Figure III.9: Bone marrow HSPCs selection and colony development. C) Merge of bright field **(A)** and fluorescence **(B)** images of bone marrow cells on the donor substrate. **D)** Culture of selected HSPCs at 1 day, 4 days **(E),** and 8 days **(F). G), H)** and **I)** depicts a close view of the colonies after 14 days. Scale bars indicate 50 μm.

Lin^- cells were printed and CFUs were observed at different times up to 14 days. Figure III.9 D, E and F shows the progression of colony formation from printed HSPCs; the HSPCs after 14 days of culture are shown in Figure III.9 G, H, and I, which shows a closer view of three distinct developed colonies. In this example, from a total of 60 HSPCs printed, 55 CFUs were detected when considering colonies of 10 or more cells, but as high as 80 when considering colonies of more than 5 cells. Therefore, the FE BA-LIFT bioprinting technique preserved 100% of the proliferation and differentiation potential of the printed HSPCs.

2. Generation of Fluorescence Reporter Cells as a Tool for the Validation of the FE BA-LIFT Bioprinting

Human *IFN-γ* genomic *locus* characterization and gRNA design

The human *IFN-γ* genomic sequence is located at chr12: 68,154,768 - 68,159,740 (UCSC Genome Browser, GRCh38/hg18 assembly) and comprises ~5 kbp with 4 coding exons. For the generation of a fluorescent reporter, the *eGFP* sequence was inserted in the C-terminal region of the IFN-γ protein, as a successful transgenic mouse had already been generated previously with the Yellow Fluorescent Protein (*YFP*) molecule in this position [251]. Human *IFN-γ* exon 4 and 3'UTR were analyzed for the presence of gRNAs by using Benchling software (M&M III.3), taking into account additional considerations when selecting the genomic target site [252]: **a)** cleavage proximity should be as close as possible to the intended integration site**; b)** the predicted on-target activity should be as high as possible**; c)** absence of exonic off-targets with high cleavage probability**;** and **d)** overlapping with the integration site must be considered**.** The latter is preferable since the target site is disrupted during template integration, thus preventing re-cleavage without having to introduce additional point mutations in the template. Selected gRNAs should cleave the DNA near the intended integration site; nevertheless, a distal target with a high on-target score may be preferred over a proximal target site with a poor on-target score.

Figure III.10 shows the quality of the predicted gRNAs targeting the region near the *IFN-γ* STOP codon, which was the intended integration site. From the predicted gRNAs analyzed using Benchling, the hIFNγ_KI_stop gRNA (Figure III.10, yellow arrow) was selected, as it met most of the abovementioned criteria.

Knock-in strategies and HDR donor design

While it is important to have an efficient and precise gRNA, it is just as important to have a well designed HDR donor that correctly integrates the sequence of interest at the intended integration site. Multiple HDR donors have been described, such as single-stranded oligodeoxyribonucleotides (ssODN) donors, single-stranded DNA

(ssDNA) fragments, plasmids, or PCR amplified double-stranded (ds) DNA donors. Choosing the most suitable donor type for introducing the intended modification depends on both the size of the modification and the selection strategy to be used latter on.

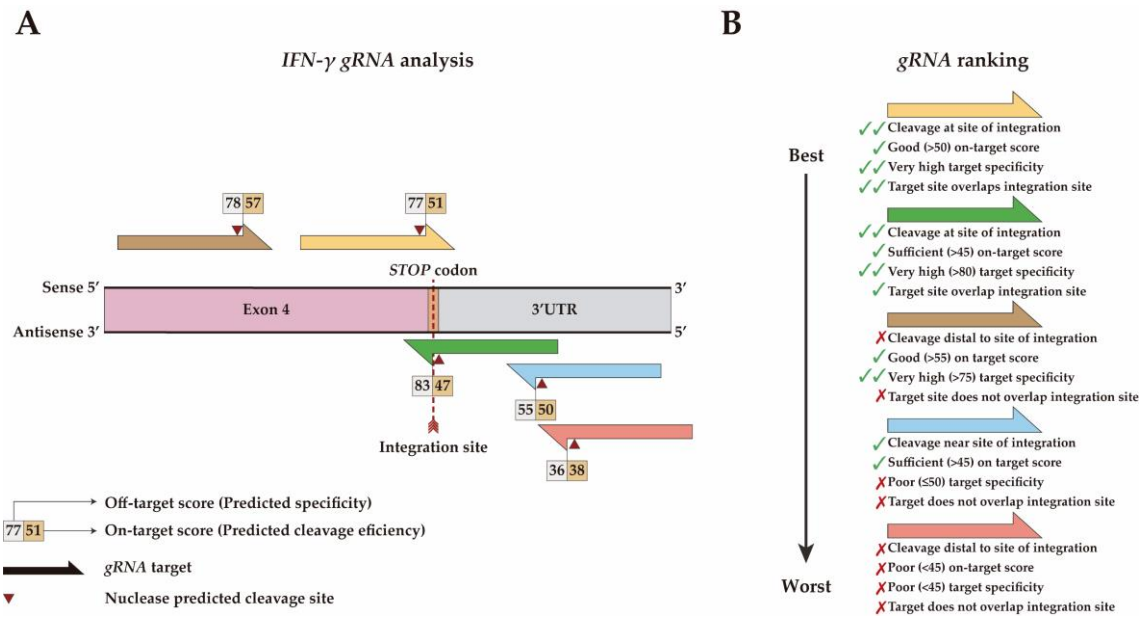


Figure III.10: Analysis and selection of a suitable gRNA for nuclease targeting. A) Schematic representation of *IFN-γ* exon 4 (dark pink), the STOP codon (orange), and the 3'UTR region (grey) with the 5-candidate gRNA (colored arrows) obtained after Benchling screening. **B)** Suitable gRNAs were ranked based on their cleavage dynamics (on- and off-target scores, the higher is better), their location and their orientation towards the intended integration site. Adapted from [252].

For the knock-in of the *eGFP* sequence after the *IFN-γ* coding sequence, two approaches were used: **1)** an in frame insertion of the *eGFP* sequence after the *IFN-γ* exon 4 (**Figure III.11**); and **2)** bicistronic expression of eGFP along with IFN-γ, by inserting the self cleaving peptide 2A (P2A), between the two coding sequences (**Figure III.12**).

The first attempt to create a fluorescent reporter cell was to create a NK-92® MI cell line expressing the IFN-γ-eGFP fusion protein by using an HDR plasmid donor. The HDR donor was generated by cloning a 2,000 bp gBlock® gene fragment into the Zero Blunt™ PCR vector (**M&M III.3**), taking special care to remove the *IFN-γ* STOP codon and changing the Leu 158 codon CTG for CTC, and Phe 159 codon TTT for TTC, to avoid the binding of the RNP complex to the HDR plasmid donor.

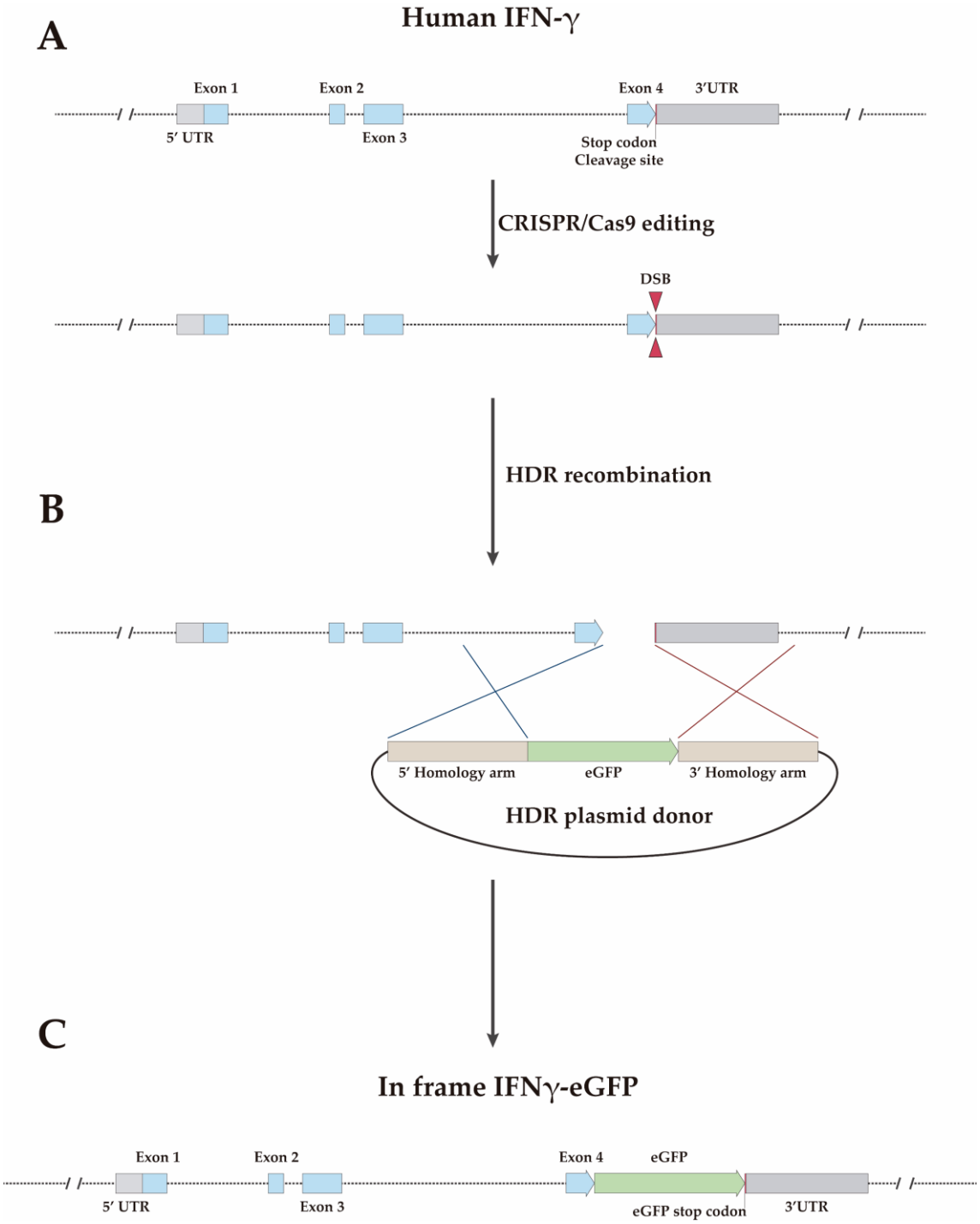


Figure III.11: Knock-in design and HDR plasmid donor strategy for the generation of IFNγ-eGFP reporter cells. A) Human *IFN-γ* gene structure and intended cleavage and insertion site location. Cleavage occurs near the STOP codon. **B)** Insertion strategy and structure of the plasmid HDR donor. **C)** Edited *IFN-γ locus* after successful integration of the *eGFP* coding sequence. An IFNγ-eGFP fusion protein will be translated.

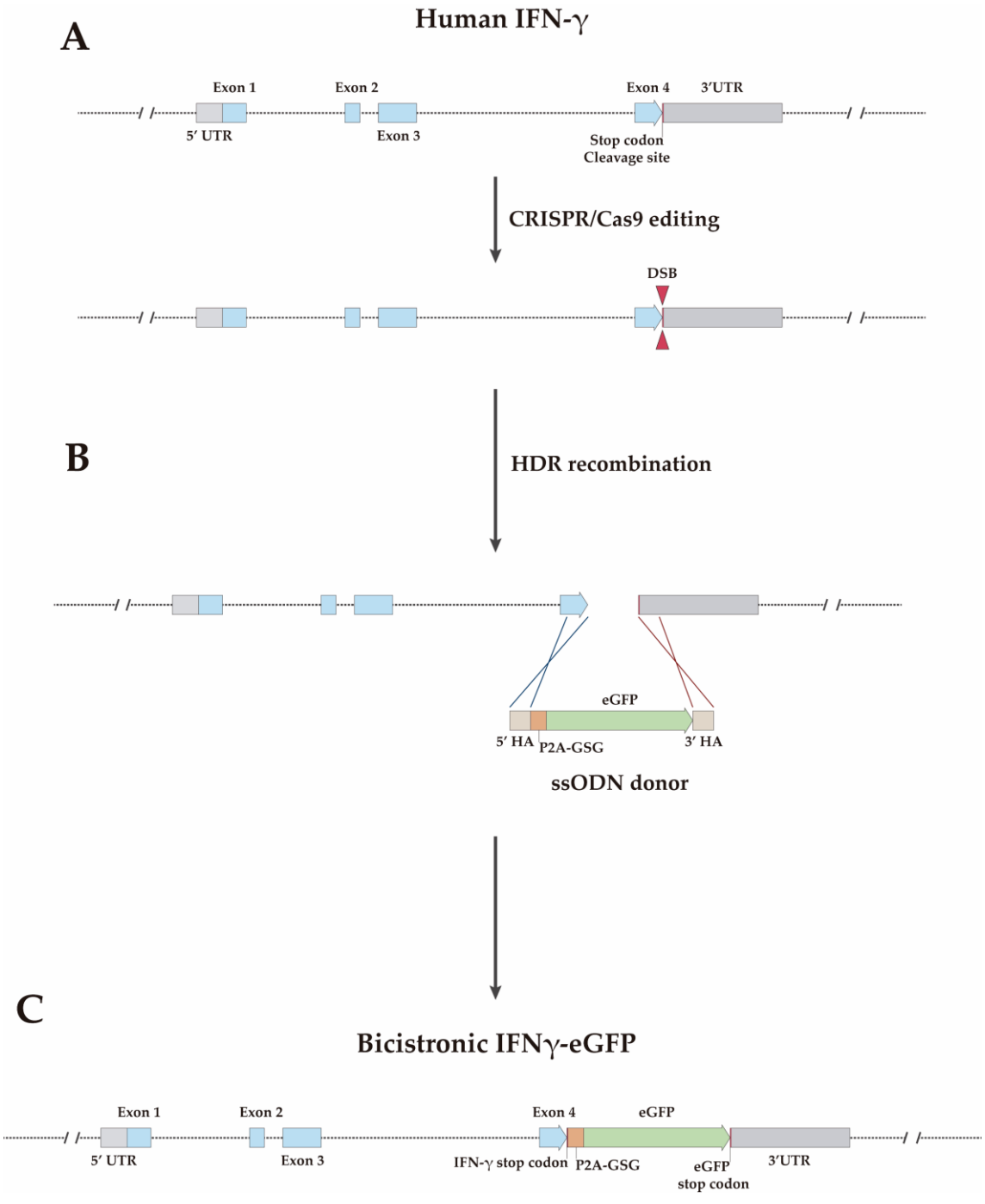


Figure III.12: Knock-in design and ssODN HDR donor strategy for the generation of IFNγ-eGFP reporter cells. A) Human *IFN-γ* gene structure and location of intended cleavage and insertion site. Cleavage will occur near the STOP codon. **B)** Insertion strategy and structure of the ssODN donor. **C)** Edited *IFN-γ locus* after successful integration of the eGFP coding sequence. The P2A-GSG peptide will be cleaved during translation, resulting in the independent expression of both proteins.

Unlike the fusion protein expression strategy, bicistronic expression ensures that any cell that is positive for the fluorescent reporter should also be expressing the gene of interest without the risk of altering its posttranslational processing. For bicistronic expression of eGFP together with IFN-γ, the coding sequence of the P2A peptide was added between the *IFN-γ* STOP codon and the *eGFP* coding sequence. The addition of the self cleaving P2A (~20 amino acids) would produce equimolecular levels of both genes from the same mRNA; theses peptides are thought to function by making the ribosome to skip the synthesis of a peptide bond at the C-terminus of the 2A element, leading to a separation between the end of the 2A sequence and the next peptide downstream [253]. Cleavage occurs between the Gly and Pro residues, meaning that the upstream cistron will have a few additional residues added to the end, while downstream cistron will start with Pro (**Figure III.13**).

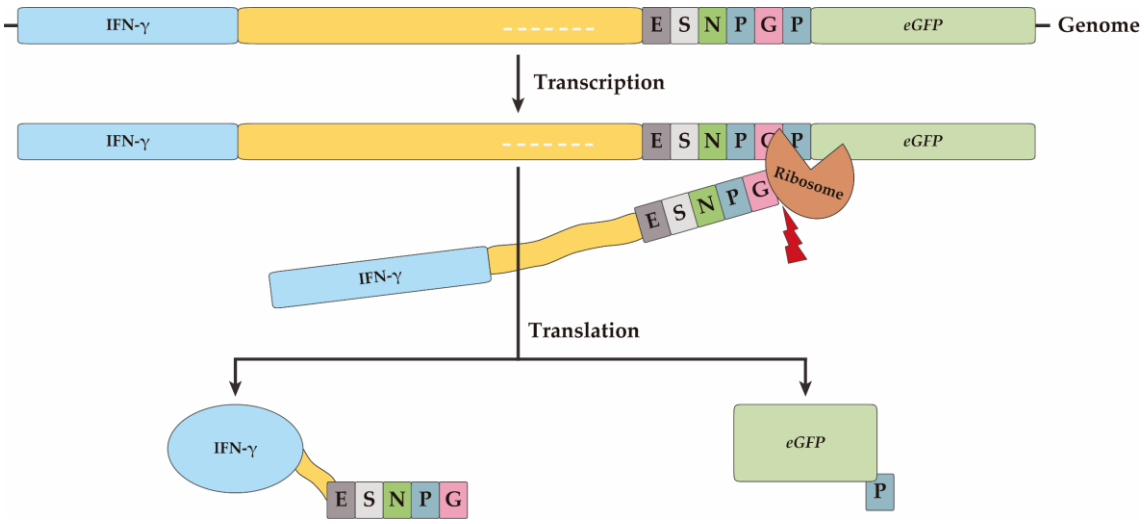


Figure III.13: Cleavage mechanism of the P2A self cleaving peptide. Cleavage site is located between Gly (G) and Pro (P) at the C-terminus of the peptide. The resulting proteins after cleavage will keep the P2A residues as follows: IFN-γ will keep the amino acids residues located at the N-terminus of the P2A; while the remaining Pro will remained attached to the eGFP protein. Adapted from [254].

Moreover, the P2A peptide with an additional GSG spacer at the N-terminus (P2A-GSG) was found to have the highest cleavage efficiency [254], so the improved P2A-GSG peptide for the bicistronic expression of IFNγ-eGFP was used. Insertion of the P2A-GSG/eGFP cassette was carried out by transfecting a single stranded oligodeoxyribonucleotide (ssODN) together with the CRISPR/Cas9 RNP complex.

Figure SIII.3 illustrate the HDR donor plasmid structure and features, as well as the sequence containing the insertion cassette, and **Figure SIII.4** illustrate the ssODN structure and sequence.

NK-92MI-eGFP genotypic and phenotypic characterization

As integration efficiencies until now have been very low (ranging from 0-50% **[255-257]**), and as the NK-92® MI cells are difficult to transfect (~30% efficiency / ~55% viability), we took advantage of the fluorescent reporter system to specifically identify and select cells with positive insertions. At 48 hours after transfection of the HDR plasmid or ssODN donors, cells were analyzed by PCR for integration of the construct (**Figure III.14 A**); positively transfected cells were then stimulated for 6 hours with 50 ng PMA and 1 μg ionomycin to allow the *IFN-γ* gene expression. Stimulated cells were then sorted with the FACSAria (Beckton Dickinson) flow cytometer, isolating and purifying the eGFP$^+$ population (**Figure III.14 B**). About 1-2% of the transfected cells were GFP$^+$, which makes the sorting a useful tool to enrich the number of edited cells, saving time on successive selection methods.

The eGFP$^+$ cells were cloned, and every clone was PCR analyzed for the integration of the eGFP construct. One homozygous clone for the insertion of the in frame donor was identified and selected (hereafter called NK-92MI *Mk. I*), whereas, for the bicistronic expression donor, only one heterozygous candidate was found (hereafter called NK-92MI *Mk. V*, **Figure III.14 A**). For both strategies, the IFNγ/eGFP protein expression was measured either in unstimulated conditions or after PMA/Io stimulation in the presence of brefeldin A, and compared by flow cytometry.

Figure III.15 shows the comparison of *eGFP* expression between the unedited NK-92® MI cells and the *Mk. I* and *Mk. V* reporter cells. Although the two reporter cells expressed eGFP after stimulation, *Mk.V* eGFP reporter was significantly more efficient (~100-fold compared to unedited cells) than the *Mk.I* fusion protein reporter (~20-fold compared with the unedited cells).

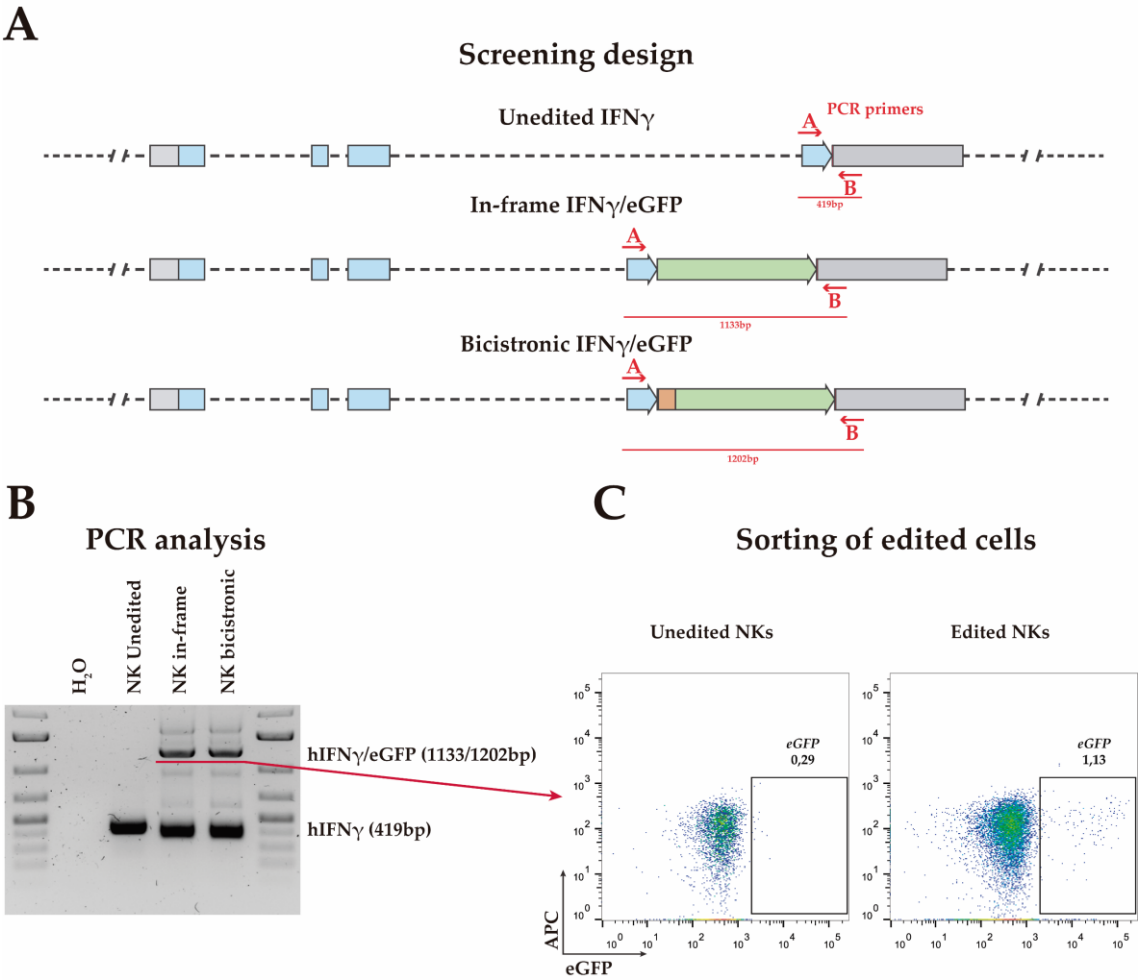


Figure III.14: Screening and sorting of NK-92® MI cells after CRISPR/Cas9 editing. A) Screening of transfected cells was carried out by PCR with primers targeting the *IFN-γ* exon 4 and 3'UTR. The expected amplicon size was 419 bp for the unedited *IFN-γ*, 1133 bp for the in frame construct insertion, and 1202 bp for the bicistronic construct insertion. **B)** Representative PCR of edited cells 48 hours after transfection**. C)** eGFP⁺ cells were sorted by FACS to discard the unedited cells. As expected, the integration efficiency was ~1-2%, so FACS represents a notable improvement in the enrichment of edited cells.

Intracytoplasmic staining (ICS) of IFN-γ showed a dramatic decrease in IFNγ-eGFP protein expression in the *Mk. I* fusion protein strategy, whereas the *Mk. V* cells expressed IFN-γ to the same extent as the unedited cells (**Figure III.16 C**). **Figure III.16 A** shows the dot plots of IFN-γ and eGFP expression. Due to the fixation during the ICS process, the eGFP was lost to some extent, making it difficult to achieve a dual staining of both markers; nevertheless, IFNγ⁺ cells in both the fusion protein and the bicistronic strategies tend to express eGFP together with the cytokine. **Figure III.16 B** shows the histogram of IFN-γ after stimulation and the decrease of cytokine expression in cells with the fusion reporter protein.

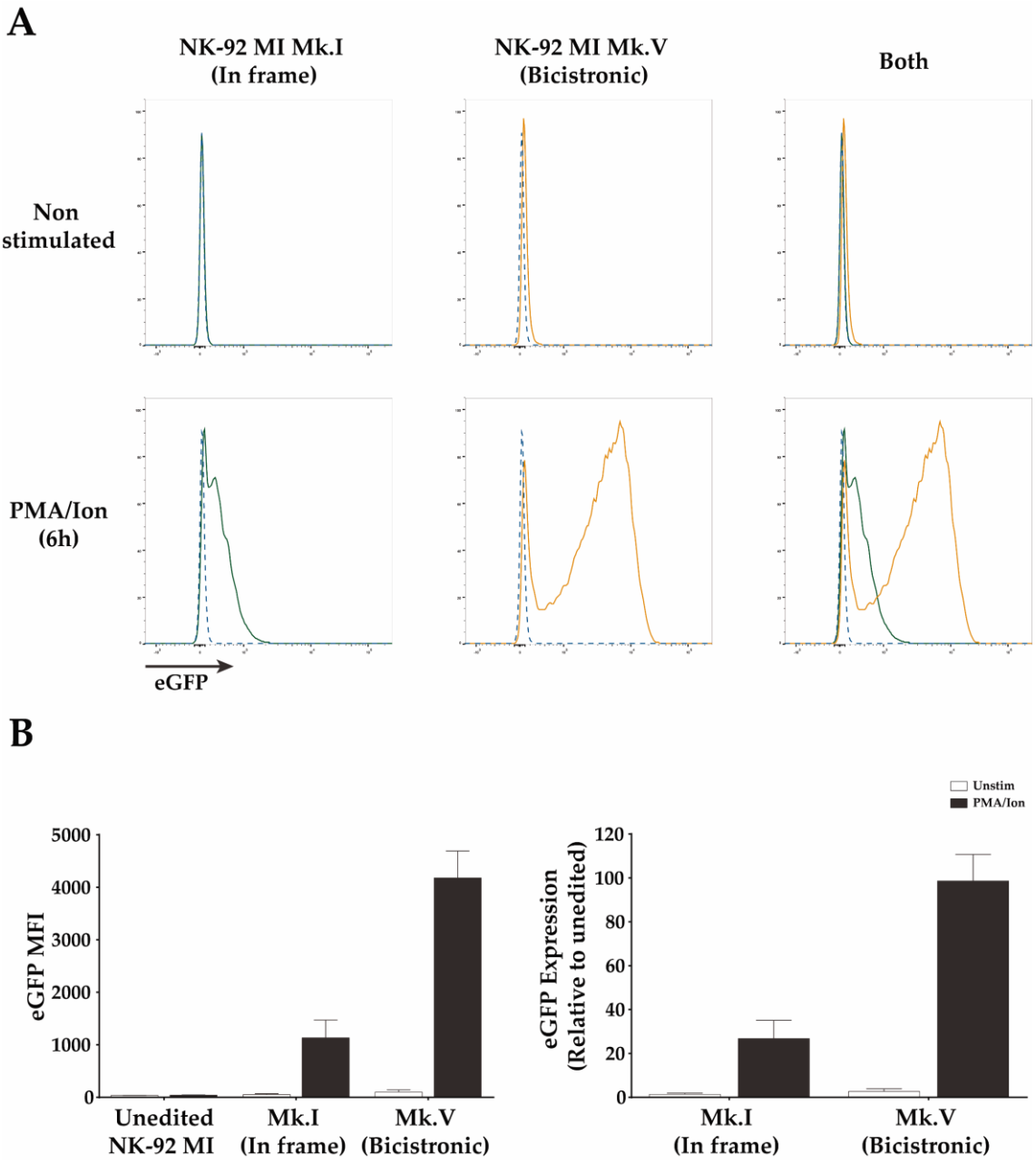


Figure III.15: Comparison of eGFP expression between the two reporter strategies shows that bicistronic expression is more efficient. A) Flow cytometry histograms comparing eGFP expression between the unedited NK-92® MI cells (blue dashed line) and NK-92 MI *Mk. I* (continuous green line), NK-92 MI *Mk. V* (continuous orange line) or both. The upper panel shows eGFP expression in unstimulated conditions, whereas the bottom panel shows eGFP expression 6 hours after PMA/Io stimulation in the presence of brefeldin A**. B)** Graphic representation of eGFP expression. The left panel shows the mean fluorescence intensity in unstimulated or stimulated condition, whereas the right panel shows the eGFP expression of the reporter cells compared to the unedited cells. Bars represents the mean ± SD of three independent experiments**. MFI:** Mean fluorescence intensity.

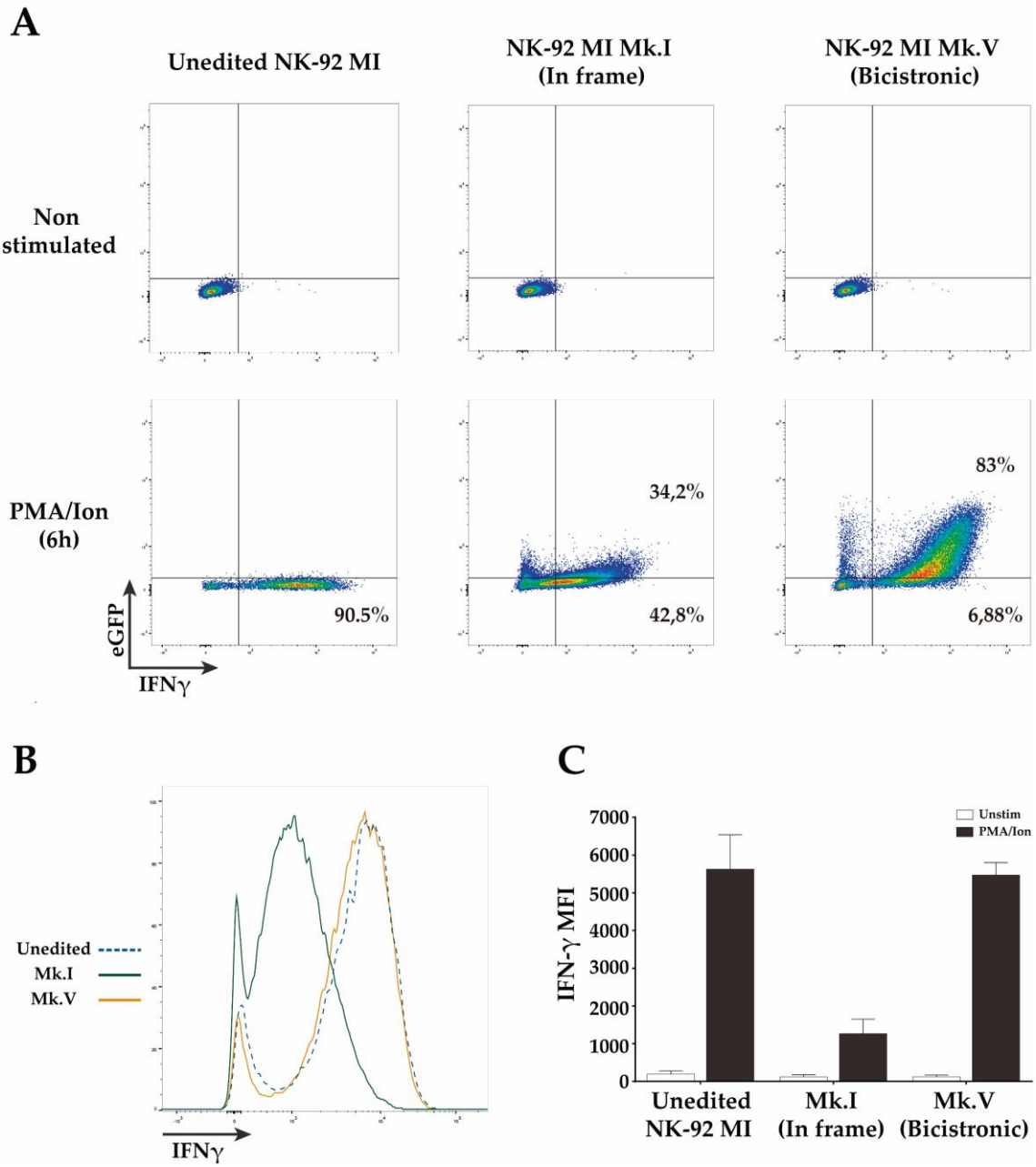


Figure III.16: Intracytoplasmic staining reveals that in frame insertion of eGFP impairs IFN-γ expression, whereas bicistronic expression does not affect cytokine expression. A) Flow cytometry dot plots comparing IFN-γ expression between the unedited NK-92® MI cells, NK-92 MI *Mk.I* and NK-92 MI *Mk.V*. The upper panel shows IFN-γ expression in unstimulated conditions, while the bottom panel shows IFN-γ expression 6 hours after PMA/Io stimulation in the presence of brefeldin A**. B)** Flow cytometry histogram comparing IFN-γ expression between the unedited NK-92® MI cells (dashed blue line), the NK-92 MI *Mk.I* (continuous green line) and the NK-92 MI *Mk.V* (continuous orange line) edited cells. **C)** Graphic representation of IFN-γ expression. Bars represent the mean ± SD of two independent experiments. **MFI:** Mean fluorescence intensity.

eGFP expression of *Mk. V* reporter cells was also characterized by confocal microscopy after stimulation with PMA/Io in the presence of brefeldin A (**Figure III.17**). CRISPR/Cas9 could successfully be applied as a tool for the generation of fluorescent reporters. It is worth mentioning that the insertion strategy needs to be optimized for every target, considering the dynamics of translational modifications of the protein of interest. Bicistronic expression has proven to be more efficient than the fusion protein approach, even though the *eGFP* cassette was inserted only in one allele of the *IFN-γ* gene. The comparison of the *INF-γ* expression between *Mk. I* and *Mk. V* cells also support the bicistronic expression as a better approach for the generation of reporter cells.

These results can potentially lead to future modifications of other immune system activating molecules, as well as cytokine genes or pro and anti inflammatory mediators.

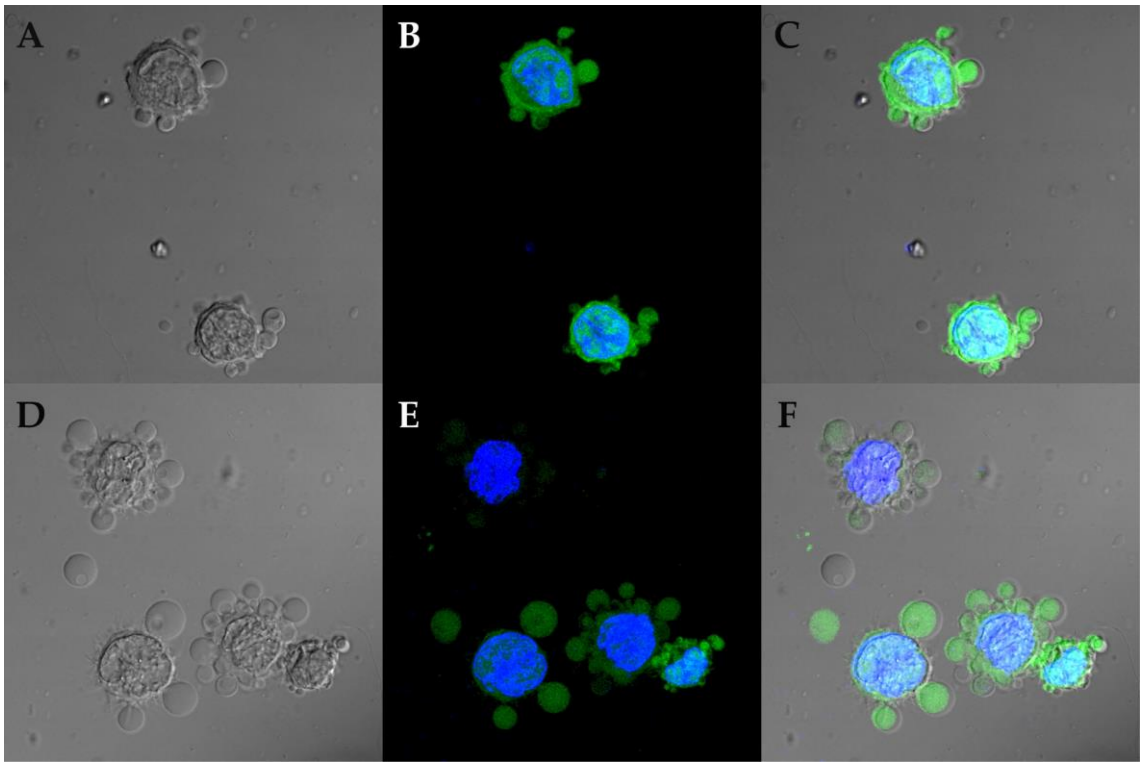


Figure III.17: eGFP expression of NK-92 *Mk. V* bicistronic reporter after PMA/Io stimulation in the presence of brefeldin A. A) and **D)** show a Normarski image of two representative fields. **B)** and **E)** show the fluorescence image of the same fields with the nucleus stained with Hoescht 33342 (blue) and the green showing eGFP expression. **C)** and **F)** correspond to the composite of both Nomarski and fluorescence images.

Discussion III

Laser bioprinting techniques as a new approach for the study of the immune system

In a new approach for the BA-LIFT technique, the fluorescence enhanced (FE) BA-LIFT has been adapted to print two different hydrogels, leading to a complete printability map. The use of an intermediate layer consisting of a commercial polyimide tape makes the process more accessible and easier to implement. The polyimide layer has two main advantages over the commonly used in LIFT: **1)** its absorption properties protect the cells from laser radiation and **2)** its transparency to visible light allows the use of an inline vision system with both conventional and fluorescence microscopy imaging, coaxial with the laser path. This set-up gives the possibility to measure cell viability before and after the printing process by applying an easy and nontoxic method to track cell survival *in situ*. Additionally, the most promising application of the FE BA-LIFT system is the possibility of selecting and printing cells previously labeled with fluorescent surface markers. These improvements make the FE BA-LIFT approach a better option than the commonly used BA-LIFT techniques [258].

Although BA-LIFT has already been performed with a polyimide layer acting as an intermediate layer [234, 258], a new approach using a commercial polyimide in the form of adhesive tape is presented. The main advantages are accessibility and the ease of manufacturing uniform polyimide coated microscope slides. Although spin coating is a very reliable process, it takes a long time to prepare each sample, and a lot of waste material is produced. Furthermore, polyimide is mechanically and chemically stable at a wide range of temperatures, ensuring that it does not degrade or release undesirable elements into the hydrogel.

To preserve the printed material and living cells, it is important to ensure that no direct laser irradiation or heating reach the biological sample [259, 260]. In the FE BA-LIFT approach, due to the very large UV absorption of the polyimide layer and its thickness (~30 μm), laser irradiation is completely absorbed within the first microns of the tape. Besides, thanks to the thermal properties of the material, there is no significant heating in the liquid layer, thus avoiding thermal effects on the biological samples.

Several studies have shown that transfer using a metallic layer is produced by complex hydro dynamical effects [259, 261]. Vaporization of the metallic layer leads to generation of a cavitation bubble which quickly expands and releases a small amount of liquid [262-264]. With this approach, some heating of the liquid is unavoidable, together with some diffusion of metallic particles in the liquid, although good cell viability can be achieved [260, 265-267].

Additionally, it was important to generate a printability map for the new FE BA-LIFT configuration for a wide range of hydrogel concentrations and set a range of parameters for the best transfer results. Previous works used a 7 μm spin coated polyimide layer, one order of magnitude thinner than the commercial tape (~30 μm). Results show that transfer is possible for a wide range of laser fluences and hydrogel viscosities, even though the polyimide thickness is significantly thicker. There have been several studies of alginate and methylcellulose laser bioprinting, mostly for sodium alginate [219, 268-271]. This part of the study aims to widen the printability study of other authors and present it as a simple summary table.

Taking advantage of the transparency of the polyimide to the visible light, a vision system was installed coaxially to the laser. Previous work has used a single standard microscopy camera to identify cells [219, 230, 231, 260, 267, 272-275]. The proposed system also integrates a fluorescence detector, allowing the system to discriminate cells not only by their shape and size but also by fluorescence markers attached to them. This is relevant for increasing the specificity of cell detection, where depending on the complexity of the fluorescence system up to 12 subpopulations of cells could be detected. This can be achieved by adding more filter sets and more light sources to the system so more fluorophores can be detected. Applications for single cell detection and analysis could use the added value of having a live and instant cell isolation system.
Additionally, by using a simple live/dead staining kit, a quick and reliable method for discriminating dead cells *in situ* from both the donor and acceptor substrates is added. During a sorting procedure all dead cells from the sample could be ignored, leaving only the viable ones for study.

The ability to trace NK cell activation at single cell level proves that cells keep their metabolic and signaling pathways. This study paves the way for future analysis of cell activation in a more complex environment like cell-cell interaction in an immunological model. The fact that printed HSPCs with the FE BA-LIFT can proliferate and form colonies validates the technique to handle very sensitive primary cell types. Moreover, given the capability to isolate and track individual cells makes the system a suitable tool for single stem cell biology analysis such as cell lineage tracking, time lapse imaging and single molecular cell profiling (proteomics, single cell genome sequencing, transcriptome sequencing and single cell RNA) [276].

Generation of fluorescence reporter cells as a tool for the validation of the FE BA-LIFT bioprinting

One way to generate a fluorescent cytokine reporter is to introduce a sequence containing the specific cytokine promoter followed by the fluorescent molecule coding region. Nevertheless, using this method has important shortcomings as the insertion occurs randomly and does not allow control of the number of copies, site of integration, or gene expression level. Additionally, chromatin dynamics and epigenetic regulators may change between cell lines, resulting in the possibility of transgene silencing, thus disabling gene reporter expression [277, 278]. Moreover, endogenous regulatory elements may interfere with the reporter promoter elements, so it is necessary to use insulator elements flanking the enhancer and its reporter gene [277, 279]. Therefore, the recently developed targeted genome editing techniques, in which a reporter gene is inserted into a specific site at the gene of interest through homologous recombination [280] after a double strand break, were used to overcome the abovementioned problems. In this approach, a genomic sequence in the form of plasmid or ssODN containing the *eGFP* sequence flanked by sequences homologous to the genomic regions at the 5′ and 3′ ends of the targeting site (which are required for HDR [281, 282]) is transfected together with the genome editing machinery.

The most common way to express a fluorescent reporter consists of a fusion protein in which the fluorescent molecule is expressed in frame with the protein of interest.

Alternatively, an internal ribosomal entry site (IRES) or viral P2A peptide may be incorporated into the targeting construct to allow the expression and translation of both the endogenous as well as the reporter genes [283]. However, it has been reported that the transcription of the second gene is less efficient than the first one when an IRES sequence is used [284, 285]. Overall, viral P2A peptide, which allows for the co-expression of multiple proteins from a single mRNA, leads to relatively high levels of downstream protein expression compared to other strategies, and due to its small size there is a lower risk of interfering with the function of co-expressed genes. P2A peptides have also been successfully employed for polycistronic and bicistronic multigene expression [254, 286-291].

In this work, a natural killer cell line was successfully genome engineered to express an IFN-γ cytokine reporter gene. Two different approaches were compared: **1)** the generation of an IFNγ-eGFP fusion protein (NK-92 MI *Mk. I* cell line), and **2)** the bicistronic expression of eGFP together with IFN-γ (NK-92 MI *Mk. V* cell line). It was found that cells with a bicistronic expression of both molecules expressed them more efficiently than those with the fusion protein expression. IFN-γ expression was reduced ~10 times in cells expressing the IFNγ-eGFP fusion protein, whereas bicistronic expression of the reporter molecules showed an expression comparable to that of the unedited NK-92 MI cells. Moreover, eGFP expression in NK-92 MI *Mk. V* cells was ~5 times higher than in NK-92 MI *Mk. I* cells, which confirms that the bicistronic expression is a much more efficient strategy for the generation of fluorescent reporters.

Supplementary Figures III

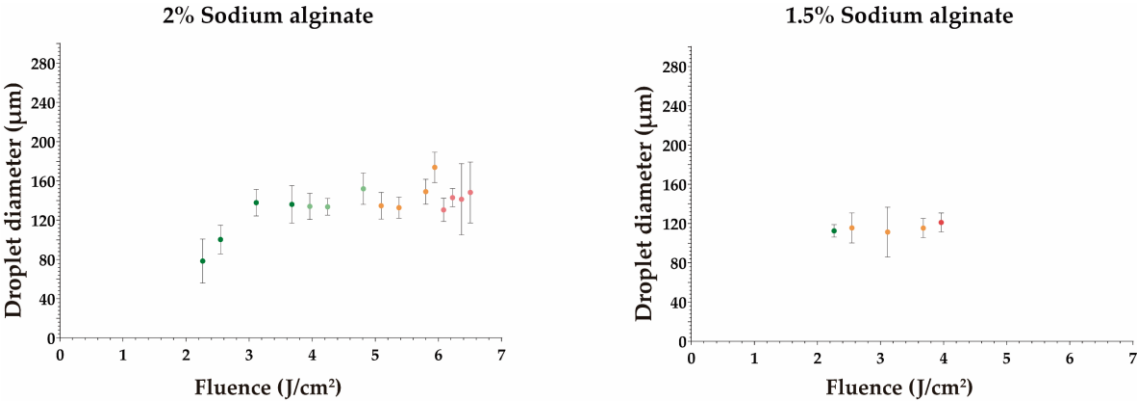


Figure SIII.1: Sodium alginate printability parametrization. Red indicates a bad shape, with a lot of satellite droplets; orange a bad shape with less satellite droplets; light green a good shape with some satellite droplets; and dark green a good shape without satellite droplets.

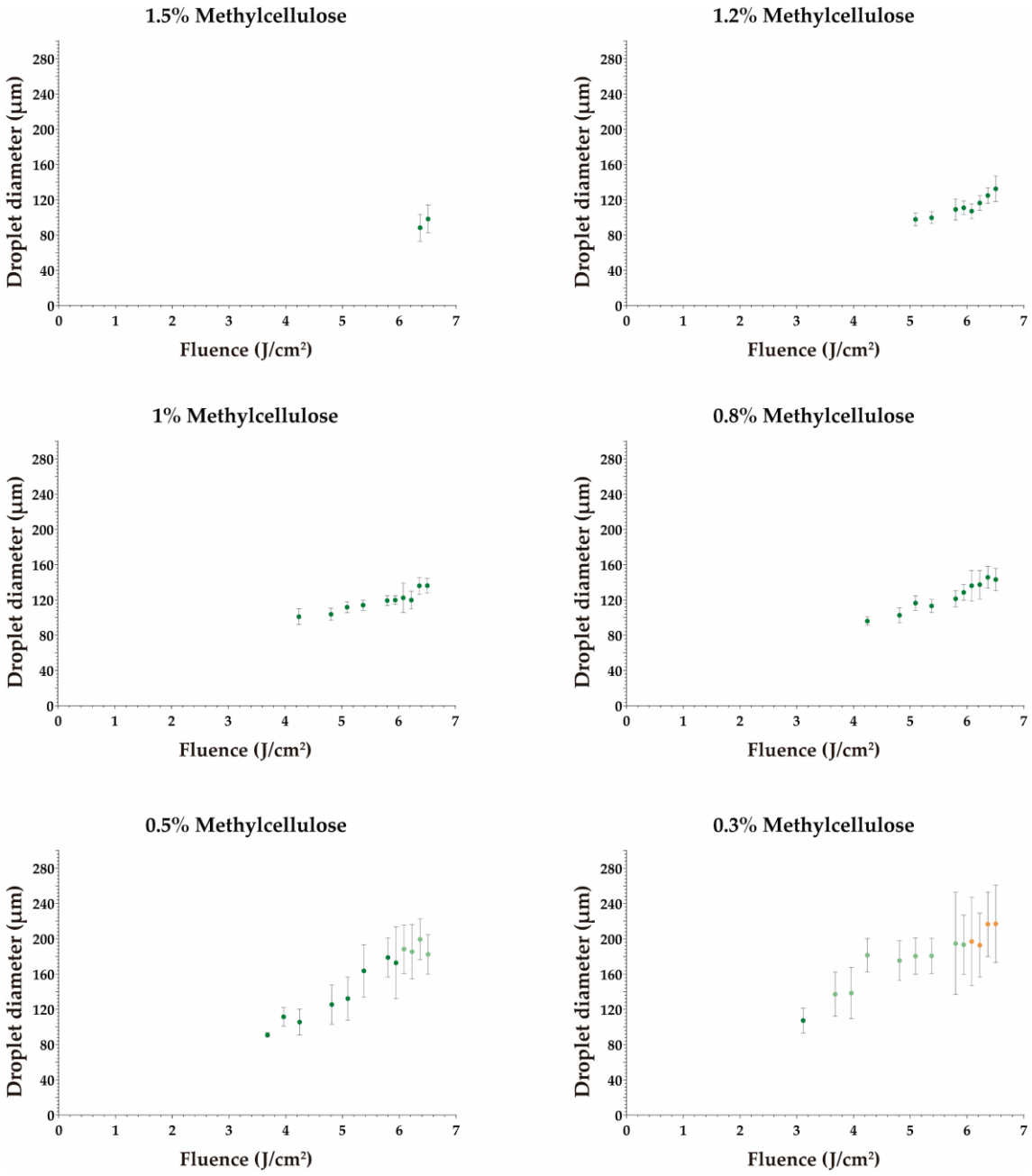
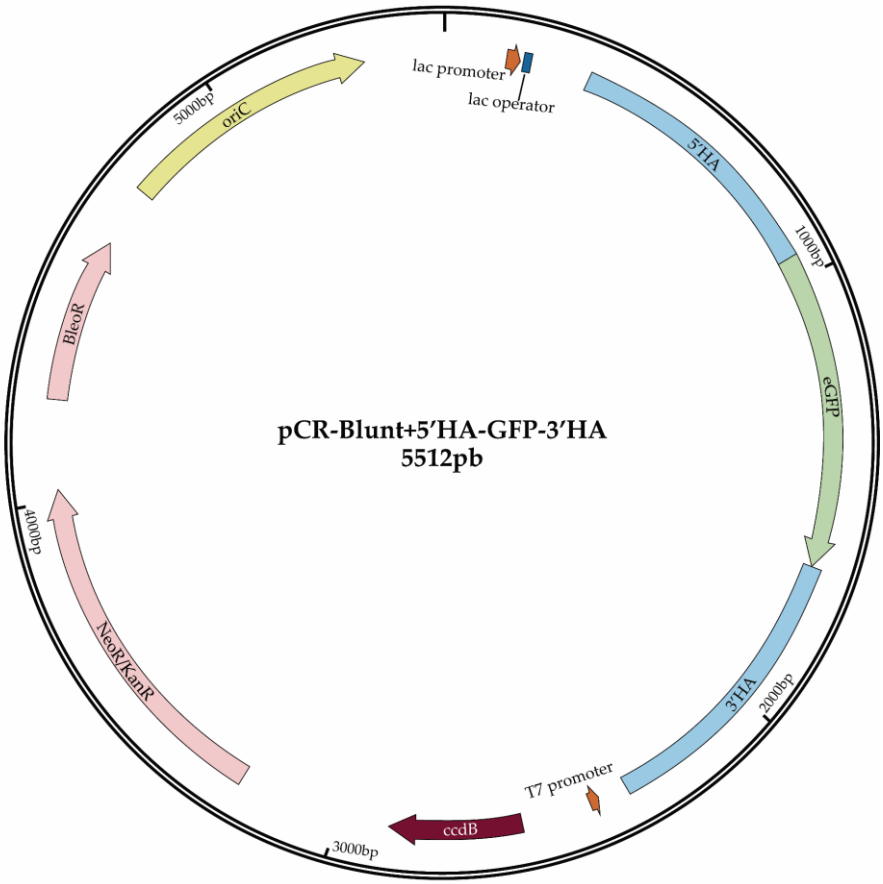


Figure SIII.2: Methylcellulose printability parametrization. Red indicates a bad shape, with a lot of satellite droplets; orange a bad shape with less satellite droplets; light green a good shape with some satellite droplets; and dark green a good shape without satellite droplets.

A



B

TGAGAATCATTGACCTATAGTTACTCATTTGATGCATACAGGAAAGACTGAAGTATAAAGTG
ATATAATTGGTAGATTGATGATAGAGAGGTCATAGAAACAGTCTCATCCTCCTTTAGATGAG
AAAATAGAAGTTCAGAGAGGTTAAGTAGCTGGCTCAAGGTCAGAATTATTGCATGCATGAG
ATTCAAACCCACCTTTTTATGCTGACTCCACAACCAGGAGTCTTTTCACTATATAATTTCAAG
AATTCTATAGAAGTAGATTTAAAGATATGTGATGGACTCCACCACATTATAGCACAACTAGA
AATGTAATTGTAATTTTTAGCTTCAACTGCTGAAGAAGTAAATATTGTATATTAAGGTAATA
CGGTCCATTTTTTAAAGGAATACTTTTATTTTCACTGACCATCATGACATTAGCAGAATATCC
TGATGGCTTATATGCCTGAAATAATTTTGCTCTTTTCTTTCCCGATAGGTAACTGACTTGAAT
GTCCAACGCAAAGCAATACATGAACTCATCCAAGTGATGGCTGAACTGTCGCCAGCAGCTA
AAACAGGGAAGCGAAAAAGGAGTCAGATGCTCTTCCGAGGTCGAAGAGCATCCCAGGTGAG
CAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGA
CCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACC
CTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTT
CAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGC
AACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGC
TGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTA
CAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTC
AAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACA
CCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCC
CTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCG
CCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAATAGTTGTCCTGCCTGCAATATTTG
AATTTTAAATCTAAATCTATTTATTAATATTTAACATTATTTATATGGGGAATATATTTTTAG
ACTCATCAATCAAATAAGTATTTATAATAGCAACTTTTGTGTAATGAAAATGAATATCTATT
AATATATGTATTATTTATAATTCCTATATCCTGTGACTGTCTCACTTAATCCTTTGTTTTCTGA
CTAATTAGGCAAGGCTATGTGATTACAAGGCTTTATCTCAGGGGCCAACTAGGCAGCCAACC
TAAGCAAGATCCCATGGGTTGTGTGTTTATTTCACTTGATGATACAATGAACACTTATAAGT
GAAGTGATACTATCCAGTTACTGCCGGTTTGAAAATATGCCTGCAATCTGAGCCAGTGCTTT
AATGGCATGTCAGACAGAACTTGAATGTGTCAGGTGACCCTGATGAAAACATAGCATCTCA
GGAGATTTCATGCCTGGTGCTTCCAAATATTGTTGACAACTGTGACTGTACCCAAATGGAAA
GTAACTCATTTGTTAAAATTATCAATATCTAATATATATGAATAAAGTGTAAGTTCACAACTACT
TATGCTGTGTTGGACTTTTTCTAAGTGAGACCTGGAGTGAAAGAACTACCTATTAATGA
ATTAGTAGGGAGGGGAGTCTTC

Figure SIII.3: Structure (A) and sequence (B) of the HDR plasmid donor. Blue letters indicate the 5′ and 3′ homology arms, and green letters the *eGFP* sequence.

A

B

GCTAAAACAGGGAAGCGAAAAAGGAGTCAGATGCTCTTCCGAGGTCGAAGAG
CATCCCAGTAAGGAAGCGGAGCTACTAACTTCAGCCTGCTGAAGCAGGCTGGC
GACGTGGAGGAGAACCCTGGACCTATGGTGAGCAAGGGCGAGGAGCTGTTCA
CCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAA
GTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCC
TGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGA
CCACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAG
CAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCAC
CATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCG
AGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGA
GGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAAC
GTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGAT
CCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAG
AACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAG
CACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTC
CTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTA
CAAGTAATAGTTGTCCTGCCTGCAATATTTGAATTTTAAATCTAAATCTATTTAT
TAATATTTAACA

Figure SIII.4: Structure (A) and sequence (B) of the HDR ssODN donor. Purple letters indicate the 5′ homology arm sequence; blue letters, the P2A-GSG sequence; green letters, the *eGFP* sequence; and grey letters the 3′ homology arm.

General Discussion

In this work, the proposed aims have been achieved by adapting the newly developed CRISPR/Cas system genome editing and the fluorescence enhanced BA-LIFT laser bioprinting tools to the study of the immune system.

CRISPR/Cas9 system tools have been used to study the CNS2 transcriptional regulation of *CD69* expression in T cells. The CRISPR-a transcriptional activation assays allowed the *in vivo* identification of stimulation responsive sequences within the CNS2 enhancer in T cells, which combined with the CRISPR-i transcriptional repression assays defined the two major areas with antagonistic but complementary regulatory activities. The deletion of the major regulatory regions within the CNS2 enhancer supports its role as a bivalent regulatory element of *CD69* expression. The relevance of these studies lies in the ability to analyze the transcription regulatory elements in their physiological conditions, which is difficult to achieve using classical approaches such as reporter assays. This is supported by the RNAseq analysis of ΔCNS2 Core E T cells, which show a dysregulation of genes related to chemokine and chemokine-receptor interactions, as previously described in mouse models.

CRISPR/Cas9 genome editing tools also demonstrate their effectiveness in generating cells lacking a functional *ERAP2* aminopeptidase in B cells in a more physiological approach than those previously described. Deletion of exons 5-6 containing the active site of the enzyme has allowed the study of its function in the class I antigen presentation pathway in the physiological context of the cell. This design is applicable to any other human cell type, allowing the generation of an *ERAP2* KO in different antigen-presenting cells.

Cytokine reporters are widely used for the study of different stages of the immune response. Nevertheless, the generation of fluorescent cytokine reporters is often a complicated process in which the insertion strategy and the resulting protein influences the expression of the target cytokine. Here, the bicistronic expression of the eGFP together with the *IFN-γ* has proven to be a more efficient approach than the generation of a fusion protein. These cells have been applied for the validation of the FE BA-LIFT bioprinting technique.

Additionally, a laser based bioprinting device for accurate identification, selection and printing of single or grouped cells within a complex population was developed and validated (FE BA-LIFT). The main advantage of our FE BA-LIFT technique is the enhancement of the system with a vision system comprising fluorescence and conventional microscopy that allows the identification and selection of cells to be printed. The vision system allows the tracking of cell survival, proliferation and activation after the printing process.

This work will serve as a basis for future technological developments and applications, representing an advance in the establishment of multidisciplinary teams that combine knowledge from areas as different as biology and engineering to develop new tools that allow us to overcome the bias of classical tools, thus expanding our knowledge regarding the behavior of the immune system cell populations. Future developments will overcome the obstacles that individual techniques from different disciplines cannot overcome.

Conclusions

Conclusions

1. By CRISPR/Cas9 toolbox assays it were defined, *in vivo*, two major areas within the CNS2 enhancer of *CD69* transcription with antagonistic but complementary regulatory activities: A Core E region with a dual repressor and activator function, and a 5′ region adjacent to the Core E (region 5′C) where the major activation transcriptional machinery is associated.

2. Chromatin immunoprecipitation suggest that Oct1 and Chd4 form part of the same complex in the CNS2 Core E, formed 24 hours after stimulation; whereas Oct1 and RUNX1 seems to be enriched coordinately outside of the Core E in the CNS2 putative poised state.

3. T cells lacking the Core E exhibit a high expression of CD69 both at basal state and upon stimulation, correlating with the enhancement of the chemokine receptors CCR1, CCR2 and CCR5 and their ligands CCL3L3, CCL4 and CCL4L1, respectively, measured by RNA-seq.

4. The effect of the loss of the human *ERAP2* aminopeptidase in the antigen presentation was assessed by generating a functional *ERAP2* KO in B cells. CRISPR/Cas9 genome editing has been successfully applied for efficiently removing of the *ERAP2* exons 5 and 6, were the active site of the protein is encoded.

5. HLA-B*27:05 ligandome analysis in ERAP2$^{-/-}$ cells determine that ERAP2 destroy ligands with N-terminal basic residues and a minority of canonical anchor motifs at P2 (lysine (K) and glutamine (Q)). A compensation effect in other peptide positions was observed, because the overall stability of the peptides remained constant.

6. FE BA-LIFT bioprinting technique is feasible for cell laden hydrogel printing at a wide range of viscosities and droplets sizes. Moreover, the enhancement of the system with a vision system comprising fluorescence and conventional microscopy allows the identification and selection of cells to be printed in the donor substrate as well as tracking cell survival, proliferation and activation after the printing process.

7. FE BA-LIFT has proven to be capable of identify and print a minor cell population from a complex population of primary cells, keeping their pluripotency, differentiation and proliferation properties after the printing process.

8. Due to the key role of IFN-γ in the immune response, we generate an eGFP reporter human NK cell (NK-92 MI) to be used in the validation of the FE BA-LIFT technique. Bicistronic expression of reporter molecules has proven to be more efficient than a fusion protein approach.

Conclusiones

Conclusiones

1. Mediante las herramientas de edición génica basadas en el sistema CRISPR/Cas9 se han definido, *in vivo*, dos áreas principales dentro de CNS2, la principal región reguladora de la transcripción de *CD69*. Estas dos áreas exhiben actividades reguladoras antagónicas pero complementarias: el núcleo transcripcional tiene una función activadora y represora; mientras que en la región 5' adyacente al núcleo (región 5'C) se asocia la principal maquinaria activadora de la transcripción de *CD69*.

2. Los experimentos de inmunoprecipitación de la cromatina sugieren que Oct1 y Chd4 forman parte del mismo complejo en el núcleo transcripcional de CNS2, formado 24 horas tras la estimulación de la célula. En estado basal, Oct1 y RUNX1 parecen enriquecerse de manera coordinada fuera del núcleo transcripcional de CNS2.

3. Las células T carentes del núcleo transcripcional de CNS2 exhiben una expresión aumentada de *CD69* tanto en reposo como tras su estimulación, lo que se correlaciona con el incremento de la expresión de los receptores de quimioquinas CCR1, CCR2 y CCR5 y sus ligandos CCL3L3, CCL4 y CCL4L1, respectivamente, analizados mediante secuenciación del ARN.

4. El efecto de la pérdida de función de la aminopeptidasa humana ERAP2 en la presentación antigénica se analizó mediante la generación de células B deficientes en la molécula ERAP2 funcional. La herramienta de edición génica CRISPR/Cas9 se aplicó satisfactoriamente para la eliminación eficiente de los exones 5 y 6 de *ERAP2*, donde se codifica el sitio activo de la proteína.

5. El análisis del ligandoma presentado por la molécula HLA-B*27:05 en células ERAP2$^{-/-}$ determina que ERAP2 destruye ligandos con residuos básicos N-terminales y una minoría de motivos de anclaje canónicos en P2 (lisina (K) y glutamina (Q)). Se aprecia un efecto compensatorio en otras posiciones del péptido, ya que la estabilidad general de los péptidos permanece constante.

6. La técnica de bioimpresión FE BA-LIFT es aplicable a la impresión de hidrogeles cargados con células en un amplio rango de viscosidades y tamaño de gota. Además, el perfeccionamiento del sistema con un sistema de visión que comprende un sistema de microscopía convencional y de fluorescencia permite la identificación y selección de las que células que se van a imprimir, así como el seguimiento y evaluación de la supervivencia, proliferación y activación celular tras el proceso de impresión.

7. La técnica de bioimpresión FE BA-LIFT ha demostrado su eficacia para la identificación e impresión de una población celular minoritaria en una mezcla compleja de células primarias, manteniendo su pluripotencia y capacidad de proliferación tras el proceso de impresión.

8. Debido al papel clave del IFN-γ en la respuesta inmunitaria, se ha generado un reportero de esta citoquina con la proteína verde fluorescente (eGFP) en células NK humanas (NK-92 MI) que ha sido posteriormente utilizada para la validación de la técnica de bioimpresión FE BA-LIFT. La expresión bicistrónica de moléculas reporteras ha demostrado ser una aproximación más eficiente que la generación de una proteína de fusión entre la citoquina y la proteína fluorescente.

References

1. Ishino, Y., et al., Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. J Bacteriol, 1987. **169**(12): p. 5429-33.
2. Barrangou, R., et al., CRISPR provides acquired resistance against viruses in prokaryotes. Science, 2007. **315**(5819): p. 1709-12.
3. Makarova, K.S., et al., An updated evolutionary classification of CRISPR-Cas systems. Nat Rev Microbiol, 2015. **13**(11): p. 722-36.
4. Makarova, K.S., et al., Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. Nat Rev Microbiol, 2019.
5. Jinek, M., et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science, 2012. **337**(6096): p. 816-21.
6. Mali, P., et al., RNA-guided human genome engineering via Cas9. Science, 2013. **339**(6121): p. 823-6.
7. Cong, L., et al., Multiplex genome engineering using CRISPR/Cas systems. Science, 2013. **339**(6121): p. 819-23.
8. Santos, R. and O. Amaral, Advances in Sphingolipidoses: CRISPR-Cas9 Editing as an Option for Modelling and Therapy. Int J Mol Sci, 2019. **20**(23).
9. Xing, H. and L.H. Meng, CRISPR-cas9: a powerful tool towards precision medicine in cancer treatment. Acta Pharmacol Sin, 2019.
10. Minkenberg, B., M. Wheatley, and Y. Yang, CRISPR/Cas9-Enabled Multiplex Genome Editing and Its Application. Prog Mol Biol Transl Sci, 2017. **149**: p. 111-132.
11. Bannikov, A.V. and A.V. Lavrov, [CRISPR/CAS9, the King of Genome Editing Tools]. Mol Biol (Mosk), 2017. **51**(4): p. 582-594.
12. Wang, H., M. La Russa, and L.S. Qi, CRISPR/Cas9 in Genome Editing and Beyond. Annu Rev Biochem, 2016. **85**: p. 227-64.
13. Zetsche, B., et al., Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. Cell, 2015. **163**(3): p. 759-71.
14. Safari, F., et al., CRISPR Cpf1 proteins: structure, function and implications for genome editing. Cell Biosci, 2019. **9**: p. 36.
15. Dai, Y., et al., Exploring the Trans-Cleavage Activity of CRISPR-Cas12a (cpf1) for the Development of a Universal Electrochemical Biosensor. Angew Chem Int Ed Engl, 2019. **58**(48): p. 17399-17405.
16. Manghwar, H., et al., CRISPR/Cas System: Recent Advances and Future Prospects for Genome Editing. Trends Plant Sci, 2019. **24**(12): p. 1102-1125.
17. Liao, H.K., et al., In Vivo Target Gene Activation via CRISPR/Cas9-Mediated Trans-epigenetic Modulation. Cell, 2017. **171**(7): p. 1495-1507 e15.
18. Simeonov, D.R., et al., Discovery of stimulation-responsive immune enhancers with CRISPR activation. Nature, 2017. **549**(7670): p. 111-115.
19. Gilbert, L.A., et al., CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. Cell, 2013. **154**(2): p. 442-51.
20. Chi, S., A. Weiss, and H. Wang, A CRISPR-Based Toolbox for Studying T Cell Signal Transduction. Biomed Res Int, 2016. **2016**: p. 5052369.
21. Anton, T., et al., Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. Nucleus, 2014. **5**(2): p. 163-72.

22. Ishii, T., et al., RNA-guided endonuclease - in situ labelling (RGEN-ISL): a fast CRISPR/Cas9-based method to label genomic sequences in various species. New Phytol, 2019. **222**(3): p. 1652-1661.
23. Chen, B., et al., Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. Cell, 2013. **155**(7): p. 1479-91.
24. Ma, H., et al., Multiplexed labeling of genomic loci with dCas9 and engineered sgRNAs using CRISPRainbow. Nat Biotechnol, 2016. **34**(5): p. 528-30.
25. Slesarev, A., et al., CRISPR/CAS9 targeted CAPTURE of mammalian genomic regions for characterization by NGS. Sci Rep, 2019. **9**(1): p. 3587.
26. Fujita, T. and H. Fujii, Efficient isolation of specific genomic regions and identification of associated proteins by engineered DNA-binding molecule-mediated chromatin immunoprecipitation (enChIP) using CRISPR. Biochem Biophys Res Commun, 2013. **439**(1): p. 132-6.
27. Fujita, T. and H. Fujii, Identification of proteins associated with an IFNgamma-responsive promoter by a retroviral expression system for enChIP using CRISPR. PLoS One, 2014. **9**(7): p. e103084.
28. Fujita, T., et al., Identification of non-coding RNAs associated with telomeres using a combination of enChIP and RNA sequencing. PLoS One, 2015. **10**(4): p. e0123387.
29. Dalvai, M., et al., A Scalable Genome-Editing-Based Approach for Mapping Multiprotein Complexes in Human Cells. Cell Rep, 2015. **13**(3): p. 621-633.
30. Savic, D., et al., CETCh-seq: CRISPR epitope tagging ChIP-seq of DNA-binding proteins. Genome Res, 2015. **25**(10): p. 1581-9.
31. Shalem, O., et al., Genome-scale CRISPR-Cas9 knockout screening in human cells. Science, 2014. **343**(6166): p. 84-7.
32. Wang, T., et al., Genetic screens in human cells using the CRISPR-Cas9 system. Science, 2014. **343**(6166): p. 80-4.
33. Higashijima, Y., et al., Applications of the CRISPR-Cas9 system in kidney research. Kidney Int, 2017. **92**(2): p. 324-335.
34. Kimura, M.Y., et al., Crucial role for CD69 in allergic inflammatory responses: CD69-Myl9 system in the pathogenesis of airway inflammation. Immunol Rev, 2017. **278**(1): p. 87-100.
35. Notario, L., et al., Anti-CD69 therapy induces rapid mobilization and high proliferation of HSPCs through S1P and mTOR. Leukemia, 2018. **32**(6): p. 1445-1457.
36. Vega-Ramos, J., et al., CD69 limits early inflammatory diseases associated with immune response to Listeria monocytogenes infection. Immunol Cell Biol, 2010. **88**(7): p. 707-15.
37. Notario, L., et al., CD69 Deficiency Enhances the Host Response to Vaccinia Virus Infection through Altered NK Cell Homeostasis. J Virol, 2016. **90**(14): p. 6464-6474.
38. Sancho, D., et al., CD69 downregulates autoimmune reactivity through active transforming growth factor-beta production in collagen-induced arthritis. J Clin Invest, 2003. **112**(6): p. 872-82.
39. Esplugues, E., et al., Enhanced antitumor immunity in mice deficient in CD69. J Exp Med, 2003. **197**(9): p. 1093-106.
40. Hara, T., et al., Human T cell activation. III. Rapid induction of a phosphorylated 28 kD/32 kD disulfide-linked early activation antigen (EA 1) by 12-o-tetradecanoyl phorbol-13-acetate, mitogens, and antigens. J Exp Med, 1986. **164**(6): p. 1988-2005.

41. Lanier, L.L., et al., Interleukin 2 activation of natural killer cells rapidly induces the expression and phosphorylation of the Leu-23 activation antigen. J Exp Med, 1988. **167**(5): p. 1572-85.
42. Alari-Pahissa, E., et al., Differential effect of CD69 targeting on bystander and antigen-specific T cell proliferation. J Leukoc Biol, 2012. **92**(1): p. 145-58.
43. Testi, R., J.H. Phillips, and L.L. Lanier, T cell activation via Leu-23 (CD69). J Immunol, 1989. **143**(4): p. 1123-8.
44. Rutella, S., et al., Induction of CD69 antigen on normal CD4+ and CD8+ lymphocyte subsets and its relationship with the phenotype of responding T-cells. Cytometry, 1999. **38**(3): p. 95-101.
45. Feng, H., et al., Listeria-infected myeloid dendritic cells produce IFN-beta, priming T cell activation. J Immunol, 2005. **175**(1): p. 421-32.
46. Shiow, L.R., et al., CD69 acts downstream of interferon-alpha/beta to inhibit S1P1 and lymphocyte egress from lymphoid organs. Nature, 2006. **440**(7083): p. 540-4.
47. Risso, A., et al., CD69 in resting and activated T lymphocytes. Its association with a GTP binding protein and biochemical requirements for its expression. J Immunol, 1991. **146**(12): p. 4105-14.
48. Bjorndahl, J.M., et al., The 28-kDa/32-kDa activation antigen EA 1. Further characterization and signal requirements for its expression. J Immunol, 1988. **141**(12): p. 4094-100.
49. Cebrian, M., et al., Expression and function of AIM, an activation inducer molecule of human lymphocytes, is dependent on the activation of protein kinase C. Eur J Immunol, 1989. **19**(5): p. 809-15.
50. Sanchez-Mateos, P., et al., Expression of a gp33/27,000 MW activation inducer molecule (AIM) on human lymphoid tissues. Induction of cell proliferation on thymocytes and B lymphocytes by anti-AIM antibodies. Immunology, 1989. **68**(1): p. 72-9.
51. Ziegler, S.F., et al., The mouse CD69 gene. Structure, expression, and mapping to the NK gene complex. J Immunol, 1994. **152**(3): p. 1228-36.
52. Gomez, M., et al., Atherosclerosis development in apolipoprotein E-null mice deficient for CD69. Cardiovasc Res, 2009. **81**(1): p. 197-205.
53. Martin, P., et al., The leukocyte activation antigen CD69 limits allergic asthma and skin contact hypersensitivity. J Allergy Clin Immunol, 2010. **126**(2): p. 355-65, 365 e1-3.
54. Esplugues, E., et al., Induction of tumor NK-cell immunity by anti-CD69 antibody therapy. Blood, 2005. **105**(11): p. 4399-406.
55. Sancho, D., et al., CD69 targeting differentially affects the course of collagen-induced arthritis. J Leukoc Biol, 2006. **80**(6): p. 1233-41.
56. Matloubian, M., et al., Lymphocyte egress from thymus and peripheral lymphoid organs is dependent on S1P receptor 1. Nature, 2004. **427**(6972): p. 355-60.
57. Radulovic, K., et al., CD69 regulates type I IFN-induced tolerogenic signals to mucosal CD4 T cells that attenuate their colitogenic potential. J Immunol, 2012. **188**(4): p. 2001-13.
58. Cruz-Adalia, A., et al., CD69 limits the severity of cardiomyopathy after autoimmune myocarditis. Circulation, 2010. **122**(14): p. 1396-404.
59. Cibrian, D. and F. Sanchez-Madrid, CD69: from activation marker to metabolic gatekeeper. Eur J Immunol, 2017. **47**(6): p. 946-953.

60. de la Fuente, H., et al., The leukocyte activation receptor CD69 controls T cell differentiation through its interaction with galectin-1. Mol Cell Biol, 2014. **34**(13): p. 2479-87.
61. Tsilingiri, K., et al., Oxidized Low-Density Lipoprotein Receptor in Lymphocytes Prevents Atherosclerosis and Predicts Subclinical Disease. Circulation, 2019. **139**(2): p. 243-255.
62. Lin, C.R., et al., Glycosylation-dependent interaction between CD69 and S100A8/S100A9 complex is required for regulatory T-cell differentiation. FASEB J, 2015. **29**(12): p. 5006-17.
63. Hayashizaki, K., et al., Myosin light chains 9 and 12 are functional ligands for CD69 that regulate airway inflammation. Sci Immunol, 2016. **1**(3): p. eaaf9154.
64. Vance, B.A., et al., Human CD69 associates with an N-terminal fragment of calreticulin at the cell surface. Arch Biochem Biophys, 2005. **438**(1): p. 11-20.
65. Cibrian, D., et al., CD69 controls the uptake of L-tryptophan through LAT1-CD98 and AhR-dependent secretion of IL-22 in psoriasis. Nat Immunol, 2016. **17**(8): p. 985-96.
66. Spiegel, S. and S. Milstien, Sphingosine-1-phosphate: an enigmatic signalling lipid. Nat Rev Mol Cell Biol, 2003. **4**(5): p. 397-407.
67. Cyster, J.G. and S.R. Schwab, Sphingosine-1-phosphate and lymphocyte egress from lymphoid organs. Annu Rev Immunol, 2012. **30**: p. 69-94.
68. Notario, L., et al., CD69 Targeting Enhances Anti-vaccinia Virus Immunity. J Virol, 2019. **93**(19).
69. Stamatoyannopoulos, J.A., What does our genome encode? Genome Res, 2012. **22**(9): p. 1602-11.
70. Lee, G.R., et al., T helper cell differentiation: regulation by cis elements and epigenetics. Immunity, 2006. **24**(4): p. 369-79.
71. Kanno, Y., et al., Transcriptional and epigenetic control of T helper cell specification: molecular mechanisms underlying commitment and plasticity. Annu Rev Immunol, 2012. **30**: p. 707-31.
72. Montavon, T. and D. Duboule, Landscapes and archipelagos: spatial organization of gene regulation in vertebrates. Trends Cell Biol, 2012. **22**(7): p. 347-54.
73. Heintzman, N.D., et al., Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet, 2007. **39**(3): p. 311-8.
74. Caglio, G., E. Torlai Triglia, and A. Pombo, Keep Them Close: PRC2 Poises Enhancer-Promoter Interactions at Anterior Neuronal Genes. Cell Stem Cell, 2017. **20**(5): p. 573-575.
75. Creyghton, M.P., et al., Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A, 2010. **107**(50): p. 21931-6.
76. Rada-Iglesias, A., et al., A unique chromatin signature uncovers early developmental enhancers in humans. Nature, 2011. **470**(7333): p. 279-83.
77. Wei, G., et al., Global mapping of H3K4me3 and H3K27me3 reveals specificity and plasticity in lineage fate determination of differentiating CD4+ T cells. Immunity, 2009. **30**(1): p. 155-67.
78. Blanco, E., et al., The Bivalent Genome: Characterization, Structure, and Regulation. Trends Genet, 2020. **36**(2): p. 118-131.

79. Rhie, S.K., et al., Identification of activated enhancers and linked transcription factors in breast, prostate, and kidney tumors by tracing enhancer networks using epigenetic traits. Epigenetics Chromatin, 2016. **9**: p. 50.
80. Lopez-Cabrera, M., et al., Molecular cloning, expression, and chromosomal localization of the human earliest lymphocyte activation antigen AIM/CD69, a new member of the C-type animal lectin superfamily of signal-transmitting receptors. J Exp Med, 1993. **178**(2): p. 537-47.
81. Lopez-Cabrera, M., et al., Transcriptional regulation of the gene encoding the human C-type lectin leukocyte receptor AIM/CD69 and functional characterization of its tumor necrosis factor-alpha-responsive elements. J Biol Chem, 1995. **270**(37): p. 21545-51.
82. Castellanos, M.C., et al., Expression of the leukocyte early activation antigen CD69 is regulated by the transcription factor AP-1. J Immunol, 1997. **159**(11): p. 5463-73.
83. Castellanos Mdel, C., et al., Multiple cis-acting elements regulate the expression of the early T cell activation antigen CD69. Eur J Immunol, 2002. **32**(11): p. 3108-17.
84. Vazquez, B.N., et al., Evidence for an intronic cis-regulatory element within CD69 gene. Genes Immun, 2012. **13**(4): p. 356-62.
85. Vazquez, B.N., et al., CD69 gene is differentially regulated in T and B cells by evolutionarily conserved promoter-distal elements. J Immunol, 2009. **183**(10): p. 6513-21.
86. Laguna, T., et al., New insights on the transcriptional regulation of CD69 gene through a potent enhancer located in the conserved non-coding sequence 2. Mol Immunol, 2015. **66**(2): p. 171-9.
87. Kent, W.J., et al., The human genome browser at UCSC. Genome Res, 2002. **12**(6): p. 996-1006.
88. Davis, C.A., et al., The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res, 2018. **46**(D1): p. D794-D801.
89. Mayor, C., et al., VISTA : visualizing global DNA sequence alignments of arbitrary length. Bioinformatics, 2000. **16**(11): p. 1046-7.
90. Dubchak, I., et al., Active conservation of noncoding sequences revealed by three-way species comparisons. Genome Res, 2000. **10**(9): p. 1304-6.
91. Frazer, K.A., et al., VISTA: computational tools for comparative genomics. Nucleic Acids Res, 2004. **32**(Web Server issue): p. W273-9.
92. Zhou, K.R., et al., ChIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data. Nucleic Acids Res, 2017. **45**(D1): p. D43-D50.
93. Cheneby, J., et al., ReMap 2018: an updated atlas of regulatory regions from an integrative analysis of DNA-binding ChIP-seq experiments. Nucleic Acids Res, 2018. **46**(D1): p. D267-D275.
94. Griffon, A., et al., Integrative analysis of public ChIP-seq experiments reveals a complex multi-cell regulatory landscape. Nucleic Acids Res, 2015. **43**(4): p. e27.
95. Oki, S., et al., ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep, 2018. **19**(12).
96. Hsu, P.D., et al., DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol, 2013. **31**(9): p. 827-32.
97. Doench, J.G., et al., Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. Nat Biotechnol, 2016. **34**(2): p. 184-191.

98. Langmead, B. and S.L. Salzberg, Fast gapped-read alignment with Bowtie 2. Nat Methods, 2012. **9**(4): p. 357-9.
99. Li, B. and C.N. Dewey, RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics, 2011. **12**: p. 323.
100. Audic, S. and J.M. Claverie, The significance of digital gene expression profiles. Genome Res, 1997. **7**(10): p. 986-95.
101. Szklarczyk, D., et al., STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res, 2019. **47**(D1): p. D607-D613.
102. Mount, D.W., Using the Basic Local Alignment Search Tool (BLAST). CSH Protoc, 2007. **2007**: p. pdb top17.
103. Altschul, S.F., et al., Basic local alignment search tool. J Mol Biol, 1990. **215**(3): p. 403-10.
104. Cheng, Y., et al., Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. Genome Res, 2009. **19**(12): p. 2172-84.
105. Tripic, T., et al., SCL and associated proteins distinguish active from repressive GATA transcription factor complexes. Blood, 2009. **113**(10): p. 2191-201.
106. Dogan, N., et al., Occupancy by key transcription factors is a more accurate predictor of enhancer activity than histone modifications or chromatin accessibility. Epigenetics Chromatin, 2015. **8**: p. 16.
107. Dittmer, J., The role of the transcription factor Ets1 in carcinoma. Semin Cancer Biol, 2015. **35**: p. 20-38.
108. Testoni, M., et al., The transcription factor ETS1 in lymphomas: friend or foe? Leuk Lymphoma, 2015. **56**(7): p. 1975-80.
109. Goswami, R. and M.H. Kaplan, STAT Transcription Factors in T Cell Control of Health and Disease. Int Rev Cell Mol Biol, 2017. **331**: p. 123-180.
110. Kumar, R. and R.A. Wang, Structure, expression and functions of MTA genes. Gene, 2016. **582**(2): p. 112-21.
111. Ma, L., et al., The Many Faces of MTA3 Protein in Normal Development and Cancers. Curr Protein Pept Sci, 2016. **17**(8): p. 726-734.
112. Muller, L., et al., NCOR1-a new player on the field of T cell development. J Leukoc Biol, 2018. **104**(6): p. 1061-1068.
113. Hrckulak, D., et al., TCF/LEF Transcription Factors: An Update from the Internet Resources. Cancers (Basel), 2016. **8**(7).
114. Mao, C.D. and S.W. Byers, Cell-context dependent TCF/LEF expression and function: alternative tales of repression, de-repression and activation potentials. Crit Rev Eukaryot Gene Expr, 2011. **21**(3): p. 207-36.
115. O'Shaughnessy, A. and B. Hendrich, CHD4 in the DNA-damage response and cell cycle progression: not so NuRDy now. Biochem Soc Trans, 2013. **41**(3): p. 777-82.
116. Williams, C.J., et al., The chromatin remodeler Mi-2beta is required for CD4 expression and T cell development. Immunity, 2004. **20**(6): p. 719-33.
117. Wang, Z., et al., Genome-wide mapping of HATs and HDACs reveals distinct functions in active and inactive genes. Cell, 2009. **138**(5): p. 1019-31.
118. Chen, M. and L.S. Qi, Repurposing CRISPR System for Transcriptional Activation. Adv Exp Med Biol, 2017. **983**: p. 147-157.

119. Dominguez, A.A., W.A. Lim, and L.S. Qi, Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. Nat Rev Mol Cell Biol, 2016. **17**(1): p. 5-15.
120. Luo, M.L., et al., Repurposing endogenous type I CRISPR-Cas systems for programmable gene repression. Nucleic Acids Res, 2015. **43**(1): p. 674-81.
121. Malina, A., et al., Repurposing CRISPR/Cas9 for in situ functional assays. Genes Dev, 2013. **27**(23): p. 2602-14.
122. Qi, L.S., et al., Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell, 2013. **152**(5): p. 1173-83.
123. Vojta, A., et al., Repurposing the CRISPR-Cas9 system for targeted DNA methylation. Nucleic Acids Res, 2016. **44**(12): p. 5615-28.
124. Shah, M., et al., Hit and Run Transcriptional Repressors Are Difficult to Catch in the Act. Bioessays, 2019. **41**(8): p. e1900041.
125. Hwang, S.S., et al., GATA-binding protein-3 regulates T helper type 2 cytokine and ifng loci through interaction with metastasis-associated protein 2. Immunology, 2010. **131**(1): p. 50-8.
126. Miccio, A. and G.A. Blobel, Role of the GATA-1/FOG-1/NuRD pathway in the expression of human beta-like globin genes. Mol Cell Biol, 2010. **30**(14): p. 3460-70.
127. Hwang, S.S., et al., Role of OCT-1 and partner proteins in T cell differentiation. Biochim Biophys Acta, 2016. **1859**(6): p. 825-31.
128. Shakya, A., et al., Oct1 and OCA-B are selectively required for CD4 memory T cell function. J Exp Med, 2015. **212**(12): p. 2115-31.
129. Shakya, A., et al., Oct1 is a switchable, bipotential stabilizer of repressed and inducible transcriptional states. J Biol Chem, 2011. **286**(1): p. 450-9.
130. Gu, X., et al., Runx1 regulation of Pu.1 corepressor/coactivator exchange identifies specific molecular targets for leukemia differentiation therapy. J Biol Chem, 2014. **289**(21): p. 14881-95.
131. Voon, D.C. and J.P. Thiery, The Emerging Roles of RUNX Transcription Factors in Epithelial-Mesenchymal Transition. Adv Exp Med Biol, 2017. **962**: p. 471-489.
132. Sander, J.D. and J.K. Joung, CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol, 2014. **32**(4): p. 347-55.
133. Gilbert, L.A., et al., Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell, 2014. **159**(3): p. 647-61.
134. Wang, Y., et al., Systematic evaluation of CRISPR-Cas systems reveals design principles for genome editing in human cells. Genome Biol, 2018. **19**(1): p. 62.
135. Wang, Y., X. Li, and H. Hu, H3K4me2 reliably defines transcription factor binding regions in different cells. Genomics, 2014. **103**(2-3): p. 222-8.
136. Zhou, L., et al., An inducible enhancer required for Il12b promoter activity in an insulated chromatin environment. Mol Cell Biol, 2007. **27**(7): p. 2698-712.
137. Fletcher, C., N. Heintz, and R.G. Roeder, Purification and characterization of OTF-1, a transcription factor regulating cell cycle expression of a human histone H2b gene. Cell, 1987. **51**(5): p. 773-81.
138. Duncliffe, K.N., et al., A T cell-specific enhancer in the interleukin-3 locus is activated cooperatively by Oct and NFAT elements within a DNase I-hypersensitive site. Immunity, 1997. **6**(2): p. 175-85.
139. Otto, F., M. Lubbert, and M. Stock, Upstream and downstream targets of RUNX proteins. J Cell Biochem, 2003. **89**(1): p. 9-18.

140. Jean-Luc Schwachtgen, J.E.R., Nathalie Janel, Reginald Brys, Danny Huylebroeck, Dominique Meyer and Danièle Kerbiriou-Nabias, Oct-1 Is Involved in the Transcriptional Repression of the von Willebrand Factor Gene Promoter. Blood, 1998. **92**(4): p. 1247-1258.
141. Kakizawa, T., et al., Silencing mediator for retinoid and thyroid hormone receptors interacts with octamer transcription factor-1 and acts as a transcriptional repressor. J Biol Chem, 2001. **276**(13): p. 9720-5.
142. dela Paz, N.G., et al., Regulation of NF-kappaB-dependent gene expression by the POU domain transcription factor Oct-1. J Biol Chem, 2007. **282**(11): p. 8424-34.
143. Inman, C.K., N. Li, and P. Shore, Oct-1 counteracts autoinhibition of Runx2 DNA binding to form a novel Runx2/Oct-1 complex on the promoter of the mammary gland-specific gene beta-casein. Mol Cell Biol, 2005. **25**(8): p. 3182-93.
144. Miccio, A., et al., NuRD mediates activating and repressive functions of GATA-1 and FOG-1 during blood development. EMBO J, 2010. **29**(2): p. 442-56.
145. Reynolds, N., et al., NuRD-mediated deacetylation of H3K27 facilitates recruitment of Polycomb Repressive Complex 2 to direct gene repression. EMBO J, 2012. **31**(3): p. 593-605.
146. González-Amaro, R., et al., Is CD69 an effective brake to control inflammatory diseases? Trends in molecular medicine, 2013. **19**(10): p. 625-632.
147. Sancho, D., et al., CD69 targeting differentially affects the course of collagen-induced arthritis. 2006. **80**(6): p. 1233-1241.
148. Shastri, N., S. Schwab, and T. Serwold, Producing nature's gene-chips: the generation of peptides for display by MHC class I molecules. Annu Rev Immunol, 2002. **20**: p. 463-93.
149. Bjorkman, P.J., et al., Structure of the human class I histocompatibility antigen, HLA-A2. J Immunol, 2005. **174**(1): p. 6-19.
150. Rammensee, H., et al., SYFPEITHI: database for MHC ligands and peptide motifs. Immunogenetics, 1999. **50**(3-4): p. 213-9.
151. Parker, K.C., M.A. Bednarek, and J.E. Coligan, Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. J Immunol, 1994. **152**(1): p. 163-75.
152. Saric, T., et al., An IFN-gamma-induced aminopeptidase in the ER, ERAP1, trims precursors to MHC class I-presented peptides. Nat Immunol, 2002. **3**(12): p. 1169-76.
153. Hattori, A., et al., Molecular cloning of adipocyte-derived leucine aminopeptidase highly related to placental leucine aminopeptidase/oxytocinase. J Biochem, 1999. **125**(5): p. 931-8.
154. Tanioka, T., et al., Human leukocyte-derived arginine aminopeptidase. The third member of the oxytocinase subfamily of aminopeptidases. J Biol Chem, 2003. **278**(34): p. 32275-83.
155. Hearn, A., I.A. York, and K.L. Rock, The specificity of trimming of MHC class I-presented peptides in the endoplasmic reticulum. J Immunol, 2009. **183**(9): p. 5526-36.
156. López de Castro, J.A., How ERAP1 and ERAP2 Shape the Peptidomes of Disease-Associated MHC-I Proteins. Frontiers in immunology, 2018. **9**: p. 2463-2463.

157. Mpakali, A., et al., Structural Basis for Antigenic Peptide Recognition and Processing by Endoplasmic Reticulum (ER) Aminopeptidase 2. J Biol Chem, 2015. **290**(43): p. 26021-32.
158. Birtley, J.R., et al., The crystal structure of human endoplasmic reticulum aminopeptidase 2 reveals the atomic basis for distinct roles in antigen processing. Biochemistry, 2012. **51**(1): p. 286-95.
159. Zervoudi, E., et al., Probing the S1 specificity pocket of the aminopeptidases that generate antigenic peptides. Biochem J, 2011. **435**(2): p. 411-20.
160. Saveanu, L., et al., Concerted peptide trimming by human ERAP1 and ERAP2 aminopeptidase complexes in the endoplasmic reticulum. Nat Immunol, 2005. **6**(7): p. 689-97.
161. York, I.A., et al., Proteolysis and class I major histocompatibility complex antigen presentation. Immunol Rev, 1999. **172**: p. 49-66.
162. Neefjes, J., et al., Towards a systems understanding of MHC class I and MHC class II antigen presentation. Nat Rev Immunol, 2011. **11**(12): p. 823-36.
163. Gonzalez-Galarza, F.F., et al., Allele frequency net 2015 update: new features for HLA epitopes, KIR and disease and HLA adverse drug reaction associations. Nucleic Acids Res, 2015. **43**(Database issue): p. D784-8.
164. Brewerton, D.A., et al., Ankylosing spondylitis and HL-A 27. Lancet, 1973. **1**(7809): p. 904-7.
165. Braun, J. and J. Sieper, Ankylosing spondylitis. Lancet, 2007. **369**(9570): p. 1379-1390.
166. Marcilla, M. and J.A. Lopez de Castro, Peptides: the cornerstone of HLA-B27 biology and pathogenetic role in spondyloarthritis. Tissue Antigens, 2008. **71**(6): p. 495-506.
167. Benjamin, R. and P. Parham, Guilt by association: HLA-B27 and ankylosing spondylitis. Immunol Today, 1990. **11**(4): p. 137-42.
168. Evans, D.M., et al., Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nature genetics, 2011. **43**(8): p. 761-767.
169. Cortes, A., et al., Major histocompatibility complex associations of ankylosing spondylitis are complex and involve further epistasis with ERAP1. Nat Commun, 2015. **6**: p. 7146.
170. Genetic Analysis of Psoriasis, C., et al., A genome-wide association study identifies new psoriasis susceptibility loci and an interaction between HLA-C and ERAP1. Nat Genet, 2010. **42**(11): p. 985-90.
171. Kirino, Y., et al., Genome-wide association analysis identifies new susceptibility loci for Behcet's disease and epistasis between HLA-B*51 and ERAP1. Nat Genet, 2013. **45**(2): p. 202-7.
172. Yin, X., et al., Genome-wide meta-analysis identifies multiple novel associations and ethnic heterogeneity of psoriasis susceptibility. Nat Commun, 2015. **6**: p. 6916.
173. Tsoi, L.C., et al., Identification of 15 new psoriasis susceptibility loci highlights the role of innate immunity. Nat Genet, 2012. **44**(12): p. 1341-8.
174. Robinson, P.C., et al., ERAP2 is associated with ankylosing spondylitis in HLA-B27-positive and HLA-B27-negative patients. Ann Rheum Dis, 2015. **74**(8): p. 1627-9.

175. International Genetics of Ankylosing Spondylitis, C., et al., Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. Nat Genet, 2013. **45**(7): p. 730-8.
176. Kuiper, J.J., et al., A genome-wide association study identifies a functional ERAP2 haplotype associated with birdshot chorioretinopathy. Hum Mol Genet, 2014. **23**(22): p. 6081-7.
177. Kuiper, J.J.W., et al., Functionally distinct ERAP1 and ERAP2 are a hallmark of HLA-A29-(Birdshot) Uveitis. Human molecular genetics, 2018. **27**(24): p. 4333-4343.
178. Tsui, F.W., et al., Association of an ERAP1 ERAP2 haplotype with familial ankylosing spondylitis. Ann Rheum Dis, 2010. **69**(4): p. 733-6.
179. Evans, D.M., et al., Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nat Genet, 2011. **43**(8): p. 761-7.
180. Wellcome Trust Case Control, C., et al., Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nat Genet, 2007. **39**(11): p. 1329-37.
181. Martin-Esteban, A., et al., Separate effects of the ankylosing spondylitis associated ERAP1 and ERAP2 aminopeptidases determine the influence of their combined phenotype on the HLA-B*27 peptidome. J Autoimmun, 2017. **79**: p. 28-38.
182. Martin-Esteban, A., et al., Functional Interaction of the Ankylosing Spondylitis-Associated Endoplasmic Reticulum Aminopeptidase 2 With the HLA-B*27 Peptidome in Human Cells. Arthritis Rheumatol, 2016. **68**(10): p. 2466-75.
183. Calvo, V., et al., Structure and diversity of HLA-B27-specific T cell epitopes. Analysis with site-directed mutants mimicking HLA-B27 subtype polymorphism. J Immunol, 1990. **144**(10): p. 4038-45.
184. Barnstable, C.J., et al., Production of monoclonal antibodies to group A erythrocytes, HLA and other human cell surface antigens-new tools for genetic analysis. Cell, 1978. **14**(1): p. 9-20.
185. Guasp, P., et al., Redundancy and Complementarity between ERAP1 and ERAP2 Revealed by their Effects on the Behcet's Disease-associated HLA-B*51 Peptidome. Mol Cell Proteomics, 2019. **18**(8): p. 1491-1510.
186. Infantes, S., et al., Multiple, non-conserved, internal viral ligands naturally presented by HLA-B27 in human respiratory syncytial virus-infected cells. Mol Cell Proteomics, 2010. **9**(7): p. 1533-9.
187. Johnstone, C., et al., The viral transcription group determines the HLA class I cellular immune response against human respiratory syncytial virus. Mol Cell Proteomics, 2015. **14**(4): p. 893-904.
188. Barnea, E., et al., The Human Leukocyte Antigen (HLA)-B27 Peptidome in Vivo, in Spondyloarthritis-susceptible HLA-B27 Transgenic Rats and the Effect of Erap1 Deletion. Mol Cell Proteomics, 2017. **16**(4): p. 642-662.
189. Alvarez-Navarro, C., et al., Endoplasmic Reticulum Aminopeptidase 1 (ERAP1) Polymorphism Relevant to Inflammatory Disease Shapes the Peptidome of the Birdshot Chorioretinopathy-Associated HLA-A*29:02 Antigen. Mol Cell Proteomics, 2015. **14**(7): p. 1770-80.

190. Cox, J. and M. Mann, MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol, 2008. **26**(12): p. 1367-72.
191. Cox, J., et al., Andromeda: a peptide search engine integrated into the MaxQuant environment. J Proteome Res, 2011. **10**(4): p. 1794-805.
192. Garcia-Medel, N., et al., Functional interaction of the ankylosing spondylitis-associated endoplasmic reticulum aminopeptidase 1 polymorphism and HLA-B27 in vivo. Mol Cell Proteomics, 2012. **11**(11): p. 1416-29.
193. Madden, D.R., et al., The three-dimensional structure of HLA-B27 at 2.1 A resolution suggests a general mechanism for tight peptide binding to MHC. Cell, 1992. **70**(6): p. 1035-48.
194. Lamas, J.R., et al., Modulation at multiple anchor positions of the peptide specificity of HLA-B27 subtypes differentially associated with ankylosing spondylitis. Arthritis Rheum, 1999. **42**(9): p. 1975-85.
195. Infantes, S., et al., Natural HLA-B*2705 protein ligands with glutamine as anchor motif: implications for HLA-B27 association with spondyloarthropathy. J Biol Chem, 2013. **288**(15): p. 10882-9.
196. Alvarez, I., et al., Identification of novel HLA-B27 ligands derived from polymorphic regions of its own or other class I molecules based on direct generation by 20 S proteasome. J Biol Chem, 2001. **276**(35): p. 32729-37.
197. Ben Dror, L., et al., The HLA-B*2705 peptidome. Arthritis Rheum, 2010. **62**(2): p. 420-9.
198. Yair-Sabag, S., et al., The Peptide Repertoire of HLA-B27 may include Ligands with Lysine at P2 Anchor Position. Proteomics, 2018. **18**(9): p. e1700249.
199. Fiorillo, M.T., et al., Susceptibility to ankylosing spondylitis correlates with the C-terminal residue of peptides presented by various HLA-B27 subtypes. Eur J Immunol, 1997. **27**(2): p. 368-73.
200. Fruci, D., et al., The peptide-binding specificity of HLA-B27 subtype (B*2705) analyzed by the use of polyalanine model peptides. Hum Immunol, 1994. **41**(1): p. 34-8.
201. Andres, A.M., et al., Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. PLoS Genet, 2010. **6**(10): p. e1001157.
202. Huang, Y., et al., 3D bioprinting and the current applications in tissue engineering. Biotechnol J, 2017. **12**(8).
203. Guillemot, F., V. Mironov, and M. Nakamura, Bioprinting is coming of age: Report from the International Conference on Bioprinting and Biofabrication in Bordeaux (3B'09). Biofabrication, 2010. **2**(1): p. 010201.
204. Mandrycky, C., et al., 3D bioprinting for engineering complex tissues. Biotechnol Adv, 2016. **34**(4): p. 422-434.
205. Hann, S.Y., et al., Recent advances in 3D printing: vascular network for tissue and organ regeneration. Transl Res, 2019. **211**: p. 46-63.
206. Loai, S., et al., Clinical Perspectives on 3D Bioprinting Paradigms for Regenerative Medicine. Regenerative Medicine Frontiers, 2019. **1**(1): p. e190004.
207. Tuan, R.S., G. Boland, and R. Tuli, Adult mesenchymal stem cells and cell-based tissue engineering. Arthritis Res Ther, 2003. **5**(1): p. 32-45.

208. Cui, X., et al., Direct human cartilage repair using three-dimensional bioprinting technology. Tissue Eng Part A, 2012. **18**(11-12): p. 1304-12.
209. Cui, X., et al., Synergistic action of fibroblast growth factor-2 and transforming growth factor-beta1 enhances bioprinted human neocartilage formation. Biotechnol Bioeng, 2012. **109**(9): p. 2357-68.
210. Cui, X., G. Gao, and Y. Qiu, Accelerated myotube formation using bioprinting technology for biosensor applications. Biotechnol Lett, 2013. **35**(3): p. 315-21.
211. Pepper, M.E., et al., Cell settling effects on a thermal inkjet bioprinter. Conf Proc IEEE Eng Med Biol Soc, 2011. **2011**: p. 3609-12.
212. Pepper, M.E., et al., Characterizing the effects of cell settling on bioprinter output. Biofabrication, 2012. **4**(1): p. 011001.
213. Gillette, B.M., et al., Dynamic hydrogels: switching of 3D microenvironments using two-component naturally derived extracellular matrices. Adv Mater, 2010. **22**(6): p. 686-91.
214. Xu, T., et al., Inkjet printing of viable mammalian cells. Biomaterials, 2005. **26**(1): p. 93-9.
215. Murphy, S.V. and A. Atala, 3D bioprinting of tissues and organs. Nat Biotechnol, 2014. **32**(8): p. 773-85.
216. Saif Khalil, W.S., Biopolymer deposition for freeform fabrication of hydrogel tissue constructs. Materials Science and Engineering: B, 2007. **27**(13).
217. Gruene, M., et al., Adipogenic differentiation of laser-printed 3D tissue grafts consisting of human adipose-derived stem cells. Biofabrication, 2011. **3**(1): p. 015005.
218. Ovsianikov, A., et al., Laser printing of cells into 3D scaffolds. Biofabrication, 2010. **2**(1): p. 014104.
219. Schiele, N.R., et al., Laser-based direct-write techniques for cell printing. Biofabrication, 2010. **2**(3): p. 032001.
220. Othon, C.M., et al., Single-cell printing to form three-dimensional lines of olfactory ensheathing cells. Biomed Mater, 2008. **3**(3): p. 034101.
221. Chen C Y, B.J.A.a.R.B.R., Cell patterning without chemical surface modification: Cell–cell interactions between printed bovine aortic endothelial cells (BAEC) on a homogeneous cell-adherent hydrogel. Applied Surface Science, 2006. **252**(24): p. 8641-8645.
222. Catros, S., et al., Layer-by-layer tissue microfabrication supports cell proliferation in vitro and in vivo. Tissue Eng Part C Methods, 2012. **18**(1): p. 62-70.
223. Michael, S., et al., Tissue engineered skin substitutes created by laser-assisted bioprinting form skin-like structures in the dorsal skin fold chamber in mice. PLoS One, 2013. **8**(3): p. e57741.
224. Barron, J.A., et al., Biological laser printing: a novel technique for creating heterogeneous 3-dimensional cell patterns. Biomed Microdevices, 2004. **6**(2): p. 139-47.
225. Koch, L., et al., Laser bioprinting of human induced pluripotent stem cells-the effect of printing and biomaterials on cell survival, pluripotency, and differentiation. Biofabrication, 2018. **10**(3): p. 035005.
226. Vinson, B.T., et al., Laser direct-write based fabrication of a spatially-defined, biomimetic construct as a potential model for breast cancer cell invasion into adipose tissue. Biofabrication, 2017. **9**(2): p. 025013.

227. Bourget, J.M., et al., Patterning of Endothelial Cells and Mesenchymal Stem Cells by Laser-Assisted Bioprinting to Study Cell Migration. Biomed Res Int, 2016. **2016**: p. 3569843.
228. Catros, S., et al., Laser-assisted bioprinting for creating on-demand patterns of human osteoprogenitor cells and nano-hydroxyapatite. Biofabrication, 2011. **3**(2): p. 025001.
229. Curley, J.L., et al., Isolated node engineering of neuronal systems using laser direct write. Biofabrication, 2016. **8**(1): p. 015013.
230. Dias, A.D., et al., Generating size-controlled embryoid bodies using laser direct-write. Biofabrication, 2014. **6**(2): p. 025007.
231. Kingsley, D.M., et al., Single-step laser-based fabrication and patterning of cell-encapsulated alginate microbeads. Biofabrication, 2013. **5**(4): p. 045006.
232. Barron, J., Application of laser printing to mammalian cells. Thin Solid Films, 2004. **453-454**: p. 383-387.
233. Turkoz, E., et al., Impulsively Induced Jets from Viscoelastic Films for High-Resolution Printing. Phys Rev Lett, 2018. **120**(7): p. 074501.
234. Brown M S, K.N.T.a.A.C.B., Time-resolved study of polyimide absorption layers for blister-actuated laser-induced forward transfer J. Appl. Phys., 2010 **107** p. 1–8.
235. Kattamis N T, P.P.E., Weiss R and Arnold C B Thick film laser induced forward transfer for deposition of thermally and mechanically sensitive materials Appl. Phys. Lett. , 2007 **91** (171120).
236. Nicholas T. Kattamis, M.S.B., Craig B. Arnold, Finite element analysis of blister formation in laser-induced forward transfer. Journal of Materials Research, 2011. **26**(18).
237. C. Frederik Brasz, J.H.Y., Craig B. Arnold, Tilting of adjacent laser-induced liquid jets. Microfluidics and Nanofluidics, 2014. **18**: p. 185-197.
238. Arnold, E.T.D.B., Comparison of jets from Newtonian and non-Newtonian fluids induced by blister-actuated laser-induced forward transfer (BA-LIFT). Applied Physics A, 2017. **123**(652).
239. Bryant, R.G., Polyimides. Ullmann's Encyclopedia of Industrial Chemistry. 2014: Wiley Online Library.
240. Starr, P., C.M. Agrawal, and S. Bailey, Biocompatibility of common polyimides with human endothelial cells for a cardiovascular microsensor. J Biomed Mater Res A, 2016. **104**(2): p. 406-12.
241. R.R.RichardsonJr, J.A.M., W.M.Reichert, Polyimides as biomaterials: preliminary biocompatibility testing. Biomaterials, 1993. **14**(8): p. 627-635.
242. Myllymaa, S., et al., Surface characterization and in vitro biocompatibility assessment of photosensitive polyimide films. Colloids Surf B Biointerfaces, 2010. **76**(2): p. 505-11.
243. Etienne Herth, K.G., Lionel Rousseau, Laurie E. Calvet & Christophe Loyez A biocompatible and flexible polyimide for wireless sensors. Microsystem Technologies, 2017. **23**: p. 5921-5929.
244. Julien, S., et al., Implantation of ultrathin, biofunctionalized polyimide membranes into the subretinal space of rats. Biomaterials, 2011. **32**(16): p. 3890-8.
245. Combs, C.A. and H. Shroff, Fluorescence Microscopy: A Concise Guide to Current Imaging Methods. Curr Protoc Neurosci, 2017. **79**: p. 2 1 1-2 1 25.

246. Chambers, D.C., et al., Transcriptomics and single-cell RNA-sequencing. Respirology, 2019. **24**(1): p. 29-36.
247. Liu, Y., et al., Advancing single-cell proteomics and metabolomics with microfluidic technologies. Analyst, 2019. **144**(3): p. 846-858.
248. Gross, A., et al., Technologies for Single-Cell Isolation. Int J Mol Sci, 2015. **16**(8): p. 16897-919.
249. Schindelin, J., et al., Fiji: an open-source platform for biological-image analysis. Nat Methods, 2012. **9**(7): p. 676-82.
250. Khalil, S. and W. Sun, Bioprinting endothelial cells with alginate for 3D tissue constructs. J Biomech Eng, 2009. **131**(11): p. 111002.
251. Reinhardt, R.L., H.E. Liang, and R.M. Locksley, Cytokine-secreting follicular T cells shape the antibody repertoire. Nat Immunol, 2009. **10**(4): p. 385-93.
252. Bollen, Y., et al., How to create state-of-the-art genetic model systems: strategies for optimal CRISPR-mediated genome editing. Nucleic Acids Res, 2018. **46**(13): p. 6435-6454.
253. Kim, J.H., et al., High cleavage efficiency of a 2A peptide derived from porcine teschovirus-1 in human cell lines, zebrafish and mice. PLoS One, 2011. **6**(4): p. e18556.
254. Wang, Y., et al., 2A self-cleaving peptide-based multi-gene expression system in the silkworm Bombyx mori. Sci Rep, 2015. **5**: p. 16273.
255. Yoshimi, K., et al., ssODN-mediated knock-in with CRISPR-Cas for large genomic regions in zygotes. Nat Commun, 2016. **7**: p. 10431.
256. Miura, H., et al., Easi-CRISPR for creating knock-in and conditional knockout mouse models using long ssDNA donors. Nat Protoc, 2018. **13**(1): p. 195-215.
257. Quadros, R.M., et al., Easi-CRISPR: a robust method for one-step generation of mice carrying conditional and insertion alleles using long ssDNA donors and CRISPR ribonucleoproteins. Genome Biol, 2017. **18**(1): p. 92.
258. Kattamis, N.T., et al., Thick film laser induced forward transfer for deposition of thermally and mechanically sensitive materials. Applied Physics Letters, 2007. **91**(17): p. 171120.
259. Unger, C., et al., Time-resolved imaging of hydrogel printing via laser-induced forward transfer. Applied Physics A, 2010. **103**(2): p. 271-277.
260. Deng, Y., et al., Single cell isolation process with laser induced forward transfer. J Biol Eng, 2017. **11**: p. 2.
261. Biver, E., et al., High-speed multi-jets printing using laser forward transfer: time-resolved study of the ejection dynamics. Opt Express, 2014. **22**(14): p. 17122-34.
262. Patrascioiu, A., et al., Laser-generated liquid microjets: correlation between bubble dynamics and liquid ejection. Microfluidics and Nanofluidics, 2013. **16**(1-2): p. 55-63.
263. Xiong R, Z.Z., Shen J, Lin Y, Huang Y and Chrisey D B Bubble Formation Modeling During Laser Direct Writing of Glycerol Solutions J. Micro Nano-Manufacturing 2014 **3** (011004).
264. Mézel, C., et al., Self-consistent modeling of jet formation process in the nanosecond laser pulse regime. Physics of Plasmas, 2009. **16**(12): p. 123112.
265. Sorkio, A., et al., Human stem cell based corneal tissue mimicking structures using laser-assisted 3D bioprinting and functional bioinks. Biomaterials, 2018. **171**: p. 57-71.

266. Ali, M., et al., Controlling laser-induced jet formation for bioprinting mesenchymal stem cells with high viability and high resolution. Biofabrication, 2014. **6**(4): p. 045001.
267. Barron, J.A., D.B. Krizman, and B.R. Ringeisen, Laser Printing of Single Cells: Statistical Analysis, Cell Viability, and Stress. Annals of Biomedical Engineering, 2005. **33**(2): p. 121-130.
268. Zhang, Z., et al., Time-Resolved Imaging Study of Jetting Dynamics during Laser Printing of Viscoelastic Alginate Solutions. Langmuir, 2015. **31**(23): p. 6447-56.
269. Yan, J., Y. Huang, and D.B. Chrisey, Laser-assisted printing of alginate long tubes and annular constructs. Biofabrication, 2013. **5**(1): p. 015002.
270. Gudapati, H., et al., Alginate gelation-induced cell death during laser-assisted cell printing. Biofabrication, 2014. **6**(3): p. 035022.
271. Fardel, R., et al., Laser forward transfer using a sacrificial layer: Influence of the material properties. Applied Surface Science, 2007. **254**(4): p. 1322-1326.
272. Riester D, Ö.A.a.W.M., Laser tool for single cell transfer J. Laser Micro Nanoeng., 2014 **9** (93): p. 7.
273. Phamduy, T.B., et al., Laser direct-write of single microbeads into spatially-ordered patterns. Biofabrication, 2012. **4**(2): p. 025006.
274. Xiong, R., et al., Study of gelatin as an effective energy absorbing layer for laser bioprinting. Biofabrication, 2017. **9**(2): p. 024103.
275. Schiele, N.R., et al., Laser direct writing of combinatorial libraries of idealized cellular constructs: Biomedical applications. Applied Surface Science, 2009. **255**(10): p. 5444-5447.
276. Etzrodt, M., M. Endele, and T. Schroeder, Quantitative single-cell approaches to stem cell research. Cell Stem Cell, 2014. **15**(5): p. 546-58.
277. Macarthur, C.C., et al., Chromatin insulator elements block transgene silencing in engineered human embryonic stem cell lines at a defined chromosome 13 locus. Stem Cells Dev, 2012. **21**(2): p. 191-205.
278. Burke, L.J. and A. Baniahmad, Co-repressors 2000. FASEB J, 2000. **14**(13): p. 1876-88.
279. Pikaart, M.J., F. Recillas-Targa, and G. Felsenfeld, Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. Genes Dev, 1998. **12**(18): p. 2852-62.
280. Davis, R.P., et al., Generation of human embryonic stem cell reporter knock-in lines by homologous recombination. Curr Protoc Stem Cell Biol, 2009. **Chapter 5**: p. Unit 5B 1 1 1-34.
281. Elliott, D.A., et al., NKX2-5eGFP/w hESCs for isolation of human cardiac progenitors and cardiomyocytes. Nature Methods, 2011. **8**(12): p. 1037-1040.
282. Den Hartogh, S.C., et al., Dual reporter MESP1 mCherry/w-NKX2-5 eGFP/w hESCs enable studying early human cardiac differentiation. Stem Cells, 2015. **33**(1): p. 56-67.
283. Trichas, G., J. Begbie, and S. Srinivas, Use of the viral 2A peptide for bicistronic expression in transgenic mice. BMC Biol, 2008. **6**: p. 40.
284. Hasegawa, K., et al., Efficient multicistronic expression of a transgene in human embryonic stem cells. Stem Cells, 2007. **25**(7): p. 1707-12.

285. Mizuguchi, H., et al., IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. Mol Ther, 2000. **1**(4): p. 376-82.
286. Geier, M., et al., Compact multi-enzyme pathways in P. pastoris. Chem Commun (Camb), 2015. **51**(9): p. 1643-6.
287. Verrier, J.D., et al., Bicistronic lentiviruses containing a viral 2A cleavage sequence reliably co-express two proteins and restore vision to an animal model of LCA1. PLoS One, 2011. **6**(5): p. e20553.
288. Tang, W., et al., Faithful expression of multiple proteins via 2A-peptide self-processing: a versatile and reliable method for manipulating brain circuits. J Neurosci, 2009. **29**(27): p. 8621-9.
289. Vaseghi, H.R., et al., Generation of an inducible fibroblast cell line for studying direct cardiac reprogramming. Genesis, 2016. **54**(7): p. 398-406.
290. Szymczak, A.L., et al., Correction of multi-gene deficiency in vivo using a single 'self-cleaving' 2A peptide-based retroviral vector. Nat Biotechnol, 2004. **22**(5): p. 589-94.
291. Takahashi, K., et al., Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell, 2007. **131**(5): p. 861-72.

Appendix

Articles

Publications, as an author or coauthor, made in the course of this thesis

Miguel G. Fontela, Laura Notario, Elisenda Alari-Pahissa, Elena Lorente and Pilar Lauzurica. 'The Conserved Non-Coding Sequence 2 (CNS2) enhances CD69 transcription through cooperation between the transcription factors Oct1 and RUNX1'. *Genes* **2019**, *10*(9), 651; https://doi.org/10.3390/genes10090651.

Andrés Márquez, **Miguel G. Fontela,** Sara Lauzurica, Rocío Candorcio-Simón, David Muñoz Martín, Miguel Morales Furio, Miguel Ubago Torres, Cristina Toledo Morcuende, Pilar Lauzurica Gómez and Carlos Molpeceres. 'Fluorescence enhanced BA-LIFT for single cell detection and isolation'. *Biofabrication* **2020**, 12-025019; https://doi.org/10.1088/1758-5090/ab6138.

Elena Lorente, **Miguel G. Fontela,** Eilon Barnea, Antonio J. Martín-Galiano, Carmen Mir, Begoña Galocha, Arie Admon, Pilar Lauzurica and Daniel Lopez. 'Modulation of natural HLA-B*27:05 ligandome by ankylosing spondylitis-associated endoplasmic reticulum aminopeptidase 2 (ERAP2)'. *Molecular and Cellular Proteomics,* **April 7, 2020,** 19 (6) 994-1004; https://doi.org/10.1074/mcp.RA120.002014.